

QY 237 EKVGRIDKNSPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTSKNTSTS 296
DB 241 EKVGRIDKNSPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTSKNTSTS 300
QY 297 RTHTEVGNAGAEVHASFEDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTAD 356
DB 301 RTHTEVGNAGAEVHASFEDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTAD 360
QY 357 TARLANIRYNTGTAPIYVNLPTTSLVLGNQTLATIKADENQISQILAPNNYPSKNL 416
DB 361 TARLANIRYNTGTAPIYVNLPTTSLVLGNQTLATIKADENQISQILAPNNYPSKNL 420
QY 417 APIALNAOKDASSTPIYNNQFLEKTKQLRLDQDVYGNIAFYNFENGRVAVDTGSN 476
DB 421 APIALNAOKDASSTPIYNNQFLEKTKQLRLDQDVYGNIAFYNFENGRVAVDTGSN 460
QY 477 MSELVPOQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEP 536
DB 461 MSELVPOQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEP 520
QY 537 NGNLOYQGGKITEFDFNFDQTSQNIKNQLAELNATNYTVLDDKIKLNAKNILIRDKRF 596
DB 521 NGNLOYQGGKITEFDFNFDQTSQNIKNQLAELNATNYTVLDDKIKLNAKNILIRDKRF 580
QY 597 HYDRNNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVI 656
DB 581 HYDRNNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVI 640
QY 657 NDRYDMLNLSLRQGGKIFDFKYNKDLPLYSNPYKVNVAVTKEINTIINSENGDT 716
DB 641 NDRYDMLNLSLRQGGKIFDFKYNKDLPLYSNPYKVNVAVTKEINTIINSENGDT 700
QY 717 STNGIKKILIFSKGYEIG 735
DB 701 STNGIKKILIFSKGYEIG 719

RESULT 9
US-09-273-839A-8
; Sequence 8, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 8
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-273-839A-8

Query Match 33.8%; Score 1275; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-84;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 486 ETARIIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPNGNIQYQOK 545
DB 29 ETARIIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPNGNIQYQOK 88
QY 546 DITEFDFNFDQTSQNIKNQLAELNATNYTVLDDKIKLNAKNILIRDKRFHYDRNNIAV 605
DB 89 DITEFDFNFDQTSQNIKNQLAELNATNYTVLDDKIKLNAKNILIRDKRFHYDRNNIAV 148
QY 606 GADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRYDMLNI 665
DB 149 GADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRYDMLNI 208

QY 666 SSLRQGGKIFDFKYNKDLPLYSNPYKVNVAVTKEINTIINSENGDTSTNGIKKIL 725
DB 209 SSLRQGGKIFDFKYNKDLPLYSNPYKVNVAVTKEINTIINSENGDTSTNGIKKIL 268
QY 726 IFSKGYEIG 735
DB 269 IFSKGYEIG 278

RESULT 10
US-08-960-780-32
; Sequence 32, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schneits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA-708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 177C8
US-08-960-780-32

Query Match 20.3%; Score 767; DB 3; Length 881;
Best Local Similarity 30.1%; Pred. No. 8.1e-47;
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;

QY 4 QENLLNESSSOGLGYFYFSDLNFOAPMVVTSSTTGDLSIPSELEN--IPSENYFQ 61
DB 39 QKNQ--QKMDRXGLLGYFYFKGKDF-SNLTMEAPTRDSTLIYDQQTANKLDDKQOEYQ 94
QY 62 SAWSGFIKKVKSDEYTFATSNADHVTWVDDQEVINKSNKIRLEKGLYQIKIQYQ 121

Db 95 SIRWGLIOSKETGDTFFNLSDDEQAIIEINGKISNKGKEQVHVHLEKGLVPKIEYQ 154
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISDNLQLPKQKSS-----N 162
Db 155 SD--TKFNIDSKTFKELFKIDSONQPOQVQOQDELNPEFNKESQBFLLAKPSKINLFT 212
QY 163 SRKQSTAGTPVDPDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSS 222
Db 213 QMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264
QY 223 PEKWTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDQSTONT 282
Db 265 PLESHTVGDPYTDYEKAARDLDSNAKETFNPLVAAPSVNVSMKVLSPENLNS--- 320
QY 283 DSETRISKNST 339
Db 321 ----NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHSET 363
QY 340 LAGERTWAETMG---LNTADTARLNANIRVYNTGTAPIYVNVLPITTSVLGKNQTLATIK 395
Db 364 VAQE--WGTSTGNTSQFTASAGYLNANVRVNVGTGAIYDKVPTTSFVL--NNDTIAIT 420
QY 396 ADENQLSQILAPNNYPSKNLAPIALNAQKASSTPTIMNVNQFLEKTKQLRLDTDOV 455
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHSHPTITLNKKQVDMLLNNKPMLETNQT 480
QY 456 YGNIAATYFNGRVRDVTGSNWEVLPIQIETTARIIFNGKOLNLVERRIAAVNPSDPLE 515
Db 481 DG---VYKIDKTHGNTVGGEMNGVQIQKAKTASIIIVDDGE--RVAEKVAAKYDENPED 536
QY 516 TTPDMLTKALKIAP--GENEFGNLQYQCKDITFDF--NFDQOTSQNIKNQLAEL-- 569
Db 537 KT-PSLTKDALKSLSPDEIKETIEGLLYKNKFIYESVMTYLDNTAKEVTKQLNDTTG 595
QY 570 ---NATNIYTVLKDILNAQWNLILIRKRPYDNRNIIAVGADSVVKEAHREVINSGTSG 626
Db 596 KFXDVSHLYDV---KLTPOKNTIK--LSILYDN---ABSNDNSIGKWTNIVSGGNG 647
QY 627 -----LLNID-----KOIRKILSYIIVEIDE-----GLKE 654
Db 648 KKQYSSNNPDANLTNTDAQELKNKRDYIISLYMKSEKNTQCEITIDGIEIPIITTKYN 707
QY 655 VINDRYDMANT--SSLRQDCKTFIDPKYNDKPLXISNPNYKVNKYAVTKENTINPSE 712
Db 708 VKNKYKRLDIIAHNIKSNPISIIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735
Db 764 IKQIYRYGKLEGDILIDKKGTHYG 790

RESULT 11
US-09-073-898-32
; Sequence 32, Application US/09073898
; Patent No. 6242669
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schneits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacy
; TITLE OF INVENTION: No. 6242669a1 Pesticidal Toxins and Nucleotide
; NUMBER OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/073,898
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 881 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: PS177C8
US-09-073-898-32

Query Match 20.3%; Score 767; DB 3; Length 881;
Best Local Similarity 30.1%; Pred. No. 8.1e-47;
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;
QY 4 QENLLNESSSSQGLLYPFDLNFQAPMVVTSSTTGDLSIPSELEN--IPSENYFQ 61
Db 39 QKNQ---QKEMDRKGLLYYFKGKDF--SNLTMPAPTRDSTLIYDQQTANKLLDKKQOEYQ 94
QY 62 SAISWGFIVKVKSEYFPAISADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIYQ 121
Db 95 SIRWGLIOSKETGDTFFNLSDDEQAIIEINGKISNKGKEQVHVHLEKGLVPKIEYQ 154
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISDNLQLPKQKSS-----N 162
Db 155 SD--TKFNIDSKTFKELFKIDSONQPOQVQOQDELNPEFNKESQBFLLAKPSKINLFT 212
QY 163 SRKQSTAGTPVDPDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSS 222
Db 213 QMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264
QY 223 PEKWTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDQSTONT 282
Db 265 PLESHTVGDPYTDYEKAARDLDSNAKETFNPLVAAPSVNVSMKVLSPENLNS--- 320
QY 283 DSETRISKNST 339
Db 321 ----NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHSET 363
QY 340 LAGERTWAETMG---LNTADTARLNANIRVYNTGTAPIYVNVLPITTSVLGKNQTLATIK 395
Db 364 VAQE--WGTSTGNTSQFTASAGYLNANVRVNVGTGAIYDKVPTTSFVL--NNDTIAIT 420
QY 396 ADENQLSQILAPNNYPSKNLAPIALNAQKASSTPTIMNVNQFLEKTKQLRLDTDOV 455
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHSHPTITLNKKQVDMLLNNKPMLETNQT 480


```
QY 456 YGNIATYNFENGRVVDTSNWSSEVLPOIQETTARIIFNGKDLNVLVERRIAAVNSDPLE 515
Db 481 DG---VYKIDTHGNIVTGEWNGVIOQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPDP 536
QY 516 TTKPDMTLKEALKIAF--GFNEPNGNLQYQCKDITEFDF--NFDQOTSQNIKNQLAEL-- 569
Db 537 KT-PSLTLDKALKSYDPDEIKEIEGLLYYKNKPIYESSVMTYLDENTAKEVTKQLNDTTG 595
QY 570 ---NATNIYTVLDKIKLNAKWNILIRDKRPHYDRNNIAGVADSVVKEAHRVINSSTEG 626
Db 596 KFKDVSHLYDV---KLTPKMNVTK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647
QY 627 -----LLNID-----KDIRKILSGYVIEIDTE-----GLKE 654
Db 648 KKQYSSNNPDANLTNTDAQEKLNKRDYIISLYMKSEKNTQCEITIDGEIYPTTKVN 707
QY 655 VINDRYDMLNI--SSLRQDGKTFIDFKYNDKULPLYISNPYKNVYAVTKENTINPSE 712
Db 708 VNKDNYKRLDIIAHNKSNPISIIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735
Db 764 IKQIYRYGKLEDDGILDKKGIHYG 790

RESULT 12
US-09-307-106-8
; Sequence 8, Application US/09307106
; Patent No. 6603063
; GENERAL INFORMATION:
; APPLICANT: Fetteison, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Finstad-Lee, George
; APPLICANT: Morrell, George
; TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/307,106
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/073,898
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
```

```
; REFERENCE/DOCKET NUMBER: MA-708C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: PS177C8a
US-09-307-106-8

Query Match 20.3%; Score 767; DB 4; Length 881;
Best Local Similarity 30.1%; Pred. No. 8.le-47;
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;

QY 4 QENLLNESSSOGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELEN--IPSENYFQ 61
Db 39 QKNQ---QKEMDRKGLLYGFKGKDF-SNLTMPAPTRDSTLIYDQQTANKLLDKKQEQY 94
QY 62 SAIWSGFIKVKSDSEYTFATSDAHNVHTWVDQEVINKASNSNKIRLEKRLYQIKYQ 121
Db 95 SIRWIGLIQSKETGDTFNLSDEQAIIIEINGKIIISNKGEKQVYHLEKGLVPIKIEYQ 154
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISSNLOLPELKOKSS-----N 162
Db 155 SD--TKFNIDSKTFKELKLFKIDSQNPQQVQODELRNPFNKESQEFLLAKPSKINLFT 212
QY 163 SRKRSTTSAGTVPDRNDGIPDSLEVEGYTVDPYKNTFLSPWISNIHKKGLTKYKSS 222
Db 213 QKMKREIDED---TDTGDSIPDLWEENGVTI-----QNRIAVKWDDSL-ASKGYTKFVSN 264
QY 223 PEKWTASDPYSDEKVTGRIDKNVSPEARHPLVAAPIVHVDMENIILSKNEDOSTQNT 282
Db 265 PLESHTVGDPDYDYEKAARDLDSNAKETENPLVAAPSVNVSMEKVLSPNENLS---- 320
QY 283 DSETRTISKNTSTSTRTHTSEVHGNAEYHASFFDIGGSVAGFSNSNSS---TVAIDHSL 339
Db 321 ---NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVMYQHS 363
QY 340 LAGERTWAETMG-----LNTADTARLNANIRVMTGTAPIYNNVLTSLVLGKQQTATIK 395
Db 364 VAQE--WGTSTGNTSQNTASAGYLNANRVNNVGTGAIYDVKPTTSFVL--NNDTIATIT 420
QY 396 ADENQLSQILAPNNYPSKNLAPIALNAQKADAGSTPITMNNQFLEKTKQLRLDQV 455
Db 421 AKSNSTALNTSPGSSYPKQGNGIAITSMDDFNSHPITLNNKQVDNLLNNKPMMLTNTQ 480
QY 456 YGNIATYNFENGRVVDTSNWSSEVLPOIQETTARIIFNGKDLNVLVERRIAAVNSDPLE 515
Db 481 DG---VYKIDTHGNIVTGEWNGVIOQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPDP 536
QY 516 TTKPDMTLKEALKIAF--GFNEPNGNLQYQCKDITEFDF--NFDQOTSQNIKNQLAEL-- 569
Db 537 KT-PSLTLDKALKSYDPDEIKEIEGLLYYKNKPIYESSVMTYLDENTAKEVTKQLNDTTG 595
QY 570 ---NATNIYTVLDKIKLNAKWNILIRDKRPHYDRNNIAGVADSVVKEAHRVINSSTEG 626
Db 596 KFKDVSHLYDV---KLTPKMNVTK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647
QY 627 -----LLNID-----KDIRKILSGYVIEIDTE-----GLKE 654
Db 648 KKQYSSNNPDANLTNTDAQEKLNKRDYIISLYMKSEKNTQCEITIDGEIYPTTKVN 707
QY 655 VINDRYDMLNI--SSLRQDGKTFIDFKYNDKULPLYISNPYKNVYAVTKENTINPSE 712
Db 708 VNKDNYKRLDIIAHNKSNPISIIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735
```

Db	764	IKQIYSRYGKLEDDGILLDKKGGIHYG	790
<p>RESULT 13</p> <p>US-09-850-351A-32</p> <p>Sequence 32, Application US/09850351A</p> <p>Patent No. 6656908</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Feitelson, Jerald S.</p> <p>Schnepf, H. Ernest</p> <p>Narva, Kenneth E.</p> <p>Stockhoff, Brian A.</p> <p>Schmeits, James</p> <p>Loewer, David</p> <p>Dullum, Charles Joseph</p> <p>Muller-Cohn, Judy</p> <p>Stamp, Lisa</p> <p>Morrill, George</p> <p>TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide Sequences Which Encode These Toxins</p> <p>NUMBER OF SEQUENCES: 144</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik</p> <p>STREET: 2421 N.W. 41st Street, Suite A-1</p> <p>CITY: Gainesville</p> <p>STATE: FL</p> <p>COUNTRY: US</p> <p>ZIP: 32606-6669</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent In Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/09/850,351A</p> <p>FILING DATE: 07-May-2001</p> <p>CLASSIFICATION: <Unknown></p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: US 09/073,898</p> <p>FILING DATE: 06-MAY-1998</p> <p>APPLICATION NUMBER: US 08/960,780</p> <p>FILING DATE: 30-OCT-1997</p> <p>APPLICATION NUMBER: US 60/029,848</p> <p>FILING DATE: 30-OCT-1996</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Sanders, Jay M.</p> <p>REGISTRATION NUMBER: 39,355</p> <p>REFERENCE/DOCKET NUMBER: MA-708CD1</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 352-375-8100</p> <p>TELEFAX: 352-372-5800</p> <p>INFORMATION FOR SEQ ID NO: 32:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 881 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: peptide</p> <p>ORIGINAL SOURCE:</p> <p>INDIVIDUAL ISOLATE: PS177C8</p> <p>SEQUENCE DESCRIPTION: SEQ ID NO: 32:</p> <p>US-09-850-351A-32</p> <p>Query Match 20.3%; Score 767; DB 4; Length 881;</p> <p>Best Local Similarity 30.1%; Pred. No. 8,1e-47;</p> <p>Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;</p>			
Qy	4	QENRLINSESSQGLLYGYSFDLNFQPMVVTSTTGDSIPSELEN--IPSENYFQ	61
Db	39	QKNQ---QKEMDRKGLLYGFGKDF-SNLTMTFAPTRDSTLIYDQQTANKLLDKKQGEYQ	94
Qy	62	SAIWSGFIKVKKSDEYTFATSNADNHTVMWVDDQEVINKASNRKLEKGRLYQIKIQYQ	121

Search completed: May 3, 2004, 19:42:47
Job time : 17.0986 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:40:13 ; Search time 36.6393 Seconds
(without alignments)
5560.545 Million cell updates/sec

Title: US-09-848-909A-13

Perfect score: 3772
Sequence: 1 EVQENLLNESSSSQGLL.....TSTNGIKLILFSKKGYEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10D_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3753	99.5	735	15	US-10-410-647-30
2	3753	99.5	736	12	US-09-848-909-1
3	3753	99.5	736	12	US-09-848-909-2
4	3753	99.5	736	12	US-09-848-909-3
5	3753	99.5	736	12	US-09-848-909-4
6	3753	99.5	736	12	US-09-848-909-5
7	3753	99.5	736	12	US-09-848-909-6
8	3753	99.5	736	12	US-09-848-909-7
9	3753	99.5	736	12	US-09-848-909-8
10	3753	99.5	736	12	US-09-848-909-9
11	3753	99.5	736	12	US-09-848-909-10
12	3753	99.5	736	12	US-09-848-909-11
13	3753	99.5	736	12	US-09-848-909-12
14	3753	99.5	736	12	US-09-848-909-13
15	3753	99.5	736	12	US-09-848-909-14

16	3753	99.5	736	12	US-09-848-909-15	Sequence 15, Appl
17	3753	99.5	736	12	US-09-848-909-16	Sequence 16, Appl
18	3753	99.5	736	12	US-09-848-909-17	Sequence 17, Appl
19	3753	99.5	736	12	US-09-848-909-18	Sequence 18, Appl
20	3753	99.5	736	12	US-09-848-909-19	Sequence 19, Appl
21	3753	99.5	736	12	US-09-848-909-20	Sequence 20, Appl
22	3753	99.5	736	12	US-09-848-909-21	Sequence 21, Appl
23	3753	99.5	736	12	US-09-848-909-22	Sequence 22, Appl
24	3753	99.5	736	12	US-10-442-502-7	Sequence 7, Appl
25	3753	99.5	763	15	US-10-442-502-5	Sequence 5, Appl
26	3753	99.5	764	15	US-10-442-502-6	Sequence 6, Appl
27	3750	99.4	735	12	US-10-402-466A-9	Sequence 9, Appl
28	3750	99.4	735	12	US-10-402-466A-13	Sequence 13, Appl
29	3750	99.4	735	14	US-10-332-282-13	Sequence 13, Appl
30	3746	99.3	735	12	US-09-848-909-30	Sequence 30, Appl
31	3746	99.3	764	12	US-10-253-286-681	Sequence 681, App
32	3746	99.3	764	15	US-10-245-871-681	Sequence 681, App
33	3646	96.7	764	9	US-09-747-521-4	Sequence 4, Appl
34	3646	96.7	764	13	US-10-106-014-4	Sequence 4, Appl
35	3646	96.7	764	13	US-10-105-695-4	Sequence 4, Appl
36	3646	96.7	764	14	US-10-105-694-7	Sequence 4, Appl
37	3042	80.6	599	12	US-09-848-909-24	Sequence 24, Appl
38	3037	80.5	595	14	US-10-332-282-11	Sequence 11, Appl
39	2910	77.1	573	12	US-10-402-466A-22	Sequence 22, Appl
40	2898	76.8	569	15	US-10-442-502-8	Sequence 8, Appl
41	2488	66.0	487	14	US-10-332-282-5	Sequence 5, Appl
42	2173	57.6	426	14	US-10-332-282-9	Sequence 9, Appl
43	2127	56.4	423	12	US-10-402-466A-24	Sequence 24, Appl
44	1624	43.1	318	14	US-10-332-282-7	Sequence 7, Appl
45	1338	35.5	258	14	US-10-332-282-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-410-647-30
; Sequence 30, Application US/10410647
; Publication No. US20030235818A1
; GENERAL INFORMATION:
; APPLICANT: PLEXUS VACCINE, INC.
; APPLICANT: Katritch, Vsevolod
; APPLICANT: Bordner, Andrew
; APPLICANT: Deans, Robert
; APPLICANT: Sumner, Mary
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES, AND METHOD OF IDENTIFYING SAME
; FILE REFERENCE: PLEX1110-1
; CURRENT APPLICATION NUMBER: US/10/410,647
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/373,668
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/371,256
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/371,250
; PRIOR FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-410-647-30

Query Match 99.5%; Score 3753; DB 15; Length 735;
Best Local Similarity 99.6%; Pred. No. 2.8e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVQENLLNESSSSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
DB 1 EVQENLLNESSSSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSGDEYTFATSDNHVTMVDQEVNKAASNKIRLEKRLYQIKIY 120

Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKEGLDFKLYWTDSONKKEV:SSDNLQLOPELQKSSNSRKKRSTASGTPVDPDRN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEV:SSDNLQLOPELQKSSNSRKKRSTASGTPVDPDRN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPVSDPEKVT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPVSDPEKVT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTISKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTISKNSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
QY 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIKADENOLSQLAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIKADENOLSQLAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNIAATYNPENGRVRVDTGSKNSEV 480
Db 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNIAATYNPENGRVRVDTGSKNSEV 480
QY 481 LPOQETIARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGL 540
Db 481 LPOQETIARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGL 540
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRFHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKVNVAATKNTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKVNVAATKNTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 2

US-09-848-909-1
; Sequence 1, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-1

Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. NO. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLLNESSESSOGLLYFSDLNFOAPMWVTSSTTGDLSIPSELENIPESENQYF 60
Db 1 EVKQENRLLNESSESSOGLLYFSDLNFOAPMWVTSSTTGDLSIPSELENIPESENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKEGLDFKLYWTDSONKKEV:SSDNLQLOPELQKSSNSRKKRSTASGTPVDPDRN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEV:SSDNLQLOPELQKSSNSRKKRSTASGTPVDPDRN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPVSDPEKVT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPVSDPEKVT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTISKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTISKNSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
QY 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIKADENOLSQLAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIKADENOLSQLAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNIAATYNPENGRVRVDTGSKNSEV 480
Db 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNIAATYNPENGRVRVDTGSKNSEV 480
QY 481 LPOQETIARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGL 540
Db 481 LPOQETIARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGL 540
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRFHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKVNVAATKNTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKVNVAATKNTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 3

US-09-848-909-2
; Sequence 2, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis

```
US-09-848-909-2
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-09-848-909-3

Query Match      99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
DB 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNSRKKRSTSGPTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNSRKKRSTSGPTVPDRDN 180
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
DB 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
DB 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
QY 541 QYQKDI TEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNAKNVILIRKRFHYDR 600
DB 541 QYQKDI TEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNAKNVILIRKRFHYDR 600
QY 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNLSLRQDGKTFIDFKKYNDKPLIYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNLSLRQDGKTFIDFKKYNDKPLIYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 4
US-09-848-909-3
; Sequence 3, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
```

```

; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-4

```

Query Match	99.5%	Score 3753;	DB 12;	Length 736;
Best Local Similarity	99.6%	Pred. No. 2.9e-274;		
Matches 732;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps	0;			

Qy	1	EVKQENRLN	ESSESSQGLGYFSD	LNFOAPMVWTSST	TDGLSPSSLENI	PSENQYF	60
Db	1	EVKQENRLN	ESSESSQGLGYFSD	LNFOAPMVWTSST	TDGLSPSSLENI	PSENQYF	60
Qy	61	QSAIWSGGI	KYVKKSDEYTF	FATSADNHVTMW	DQEVINKASNSKIRLEKGLRYQIKQY	120	
Db	61	QSAIWSGGI	KYVKKSDEYTF	FATSADNHVTMW	DQEVINKASNSKIRLEKGLRYQIKQY	120	
Qy	121	QRENTEKGLD	FKLYWTDSONKKEVI	SSDNLQIPELKQKSNSRKKGR	STSAGTVPDRDN	180	
Db	121	QRENTEKGLD	FKLYWTDSONKKEVI	SSDNLQIPELKQKSNSRKKGR	STSAGTVPDRDN	180	
Qy	181	DGIPDLSVEG	YTVDVKNKRTFLSP	WISNIHEKKGLTKYKSSPEKWS	TASDYSDFEKT	240	
Db	181	DGIPDLSVEG	YTVDVKNKRTFLSP	WISNIHEKKGLTKYKSSPEKWS	TASDYSDFEKT	240	
Qy	241	GRIDKNVSPEAR	HPHVAAYPIVHV	DMENILSKNEQOSTONTDSE	TRTTSKNTSRTHT	300	
Db	241	GRIDKNVSPEAR	HPHVAAYPIVHV	DMENILSKNEQOSTONTDSE	TRTTSKNTSRTHT	300	
Qy	301	SEVHGNAEV	HASFFDIGGSV	SAGFSNSNSTVAIDHSL	SLAGERTWAE	TWGLNTADTARL	360
Db	301	SEVHGNAEV	HASFFDIGGSV	SAGFSNSNSTVAIDHSL	SLAGERTWAE	TWGLNTADTARL	360
Qy	361	NANIRYVNTG	TAPIYVNVPTTSL	VLGNKOTLATIKADENQL	SOILAPNRYPSKNLAPIA	420	
Db	361	NANIRYVNTG	TAPIYVNVPTTSL	VLGNKOTLATIKAKENQL	SOILAPNRYPSKNLAPIA	420	
Qy	421	LNAQKDA	SSTPTIMYNGOFLE	EKT'KOLRLD	TDQVGNIAFYFENGVRV	VDTGSNWSEV	480
Db	421	LNAQDFS	STPTIMYNGOFLE	EKT'KOLRLD	TDQVGNIAFYFENGVRV	VDTGSNWSEV	480
Qy	481	LPQIETTARII	FNGKDLN	VERRIAANVNSDP	LETTKPDMTLKEAL	KIAPGFNPNGNL	540
Db	481	LPQIETTARII	FNGKDLN	VERRIAANVNSDP	LETTKPDMTLKEALKI	APGFNPNGNL	540
Qy	541	QYCGKDI	TEFDNFPQOTS	SONIKQLAELNATNI	YTVLDKI	KLNAKWNILIRDKRPHYDR	600
Db	541	QYCGKDI	TEFDNFPQOTS	SONIKQLAELNATNI	YTVLDKI	KLNAKWNILIRDKRPHYDR	600
Qy	601	NNIAGADESV	YKEAHRVINS	TSEGLLLNTDKDIRKILSGY	IVBIETEG	LKEVINDRY	660
Db	601	NNIAGADESV	YKEAHRVINS	TSEGLLLNTDKDIRKILSGY	IVBIETEG	LKEVINDRY	660
Qy	661	DLNLTSSLRQ	DGKTIIDFKYND	KLPLYSNPNYKVN	YAVTKENTI	INPSENGDTSTNG	720
Db	661	DLNLTSSLRQ	DGKTIIDFKYND	KLPLYSNPNYKVN	YAVTKENTI	INPSENGDTSTNG	720
Qy	721	IKKILFSK	KGYEIG	735			
Db	721	IKKILFSK	KGYEIG	735			

RESULT 6
US-09-848-909-5

```

: Sequence 5, Application US/09848909
: Publication No. US20020039588A1
: GENERAL INFORMATION:
: APPLICANT: Collier, R. John
: APPLICANT: Sellman, Brett R.
: TITLE OF INVENTION: Compounds and Methods for the Treatment
: TITLE OF INVENTION: and prevention of Bacterial Infection
: FILE REFERENCE: 00742/060002
: CURRENT APPLICATION NUMBER: US/09/848,909
: CURRENT FILING DATE: 2001-05-04
: PRIOR APPLICATION NUMBER: US 60/201,800
: PRIOR FILING DATE: 2000-04-04
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 736
: TYPE: PRT
: ORGANISM: Bacillus anthracis
: US-09-848-909-5

```

```

Query Match          99.5%;  Score 3753;  DB 12;  Length 736;
Best Local Similarity 99.6%;  Pred. No. 2.9e-274;

```

Qy	1	EVKQENRLNNESSQGLLYGYFSDLNQAPMWVTSSITGDLSTPSSLENIIPSENQYF	60
Db	1	EVKQENRLNNESSQGLLYGYFSDLNQAPMWVTSSITGDLSTPSSLENIIPSENQYF	60
Qy	61	QSAINSGFTKVKKSDEYTFATGADNHVTMWDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
Db	61	QSAINSGFTKVKKSDEYTFATGADNHVTMWDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
Qy	121	QRENPTKEGLDFKLYWYDSONKKEVISSNLQULPELKQKSSNSRKRSSTAGFTVPDRDN	180
Db	121	QRENPTKEGLDFKLYWYDSONKKEVISSNLQULPELKQKSSNSRKRSSTAGFTVPDRDN	180
Qy	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKXWSTASDPYSDPEKVT	240
Db	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKXWSTASDPYSDPEKVT	240
Qy	241	GRIDKNVSPEARHPLVAAPIYVHVDMENIILSKNEDQSTQNTDSETRTTSKNTSTSRTH	300
Db	241	GRIDKNVSPEARHPLVAAPIYVHVDMENIILSKNEDQSTQNTDSETRTTSKNTSTSRTH	300
Qy	301	SEVHGNAEVAHSPFDLGGVSAGFSNSNSTVAIDHSLSLAGERTWAEFTWGLNTADTARL	360
Db	301	SEVHGNAEVAHSPFDLGGVSAGFSNSNSTVAIDHSLSLAGERTWAEFTWGLNTADTARL	360
Qy	361	NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSOILAPNNYPSKNLAPIA	420
Db	361	NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSOILAPNNYPSKNLAPIA	420
Qy	421	LNAQKASSTPTIMVYNQFLELEKTKQLRLDQVYGNATYVNFENGVRVDTGSKWSEV	480
Db	421	LNAQDFSPTPTIMVYNQFLELEKTKQLRLDQVYGNATYVNFENGVRVDTGSKWSEV	480
Qy	481	LPQIQETTARIIFNGKDLNLVERRIIAVNPSPDLETTPKPDMLTKEALKIAPGNEPNGNL	540
Db	481	LPQIQETTARIIFNGKDLNLVERRIIAVNPSPDLETTPKPDMLTKEALKIAPGNEPNGNL	540
Qy	541	OYQGGKITEPDPNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKNMILIRDKRHHYDR	600
Db	541	OYQGGKITEPDPNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKNMILIRDKRHHYDR	600
Qy	601	NNIYAVGADESVMKEAAREVINSSTEGILLNIDKDIRKILSGYIVEIETDGLKEVINDRY	660
Db	601	NNIYAVGADESVMKEAAREVINSSTEGILLNIDKDIRKILSGYIVEIETDGLKEVINDRY	660
Qy	661	DMLNISLRODGKTFIDPFKYNDKLPYISNPNYKNVYAVTKENIILINPSENGDSTNG	720
Db	661	DMLNISLRODGKTFIDPFKYNDKLPYISNPNYKNVYAVTKENIILINPSENGDSTNG	720
Qy	721	IKKILIPFSKKGYBIG	735

Db 721 IKKILIFSKGYEIG 735
RESULT 7
US-09-848-909-6
; Sequence 6, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 06742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-6
Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIVKKSDEVTATSDNHNVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIVKKSDEVTATSDNHNVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQJPELKOKSSNSRKKRSTASGTPVDPDRN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQJPELKOKSSNSRKKRSTASGTPVDPDRN 180
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASPFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASPFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYNNVLPITSLVGNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYNNVLPITSLVGNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTITWNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKDASSPTITWNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKAFGNEPENG 540
QY 541 QYQKDI TEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAKONILIRDRKPHYDR 600
Db 541 QYQKDI TEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAKONILIRDRKPHYDR 600
QY 601 NNTIAGDAESVWKEAHEVINSSTEGILLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNTIAGDAESVWKEAHEVINSSTEGILLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660

QY 561 DMLNIISSIRODGKTFIDFKYNDKLPYISNPNYKVNYYAVTKENTIIINSENGDISTNG 720
Db 561 DMLNIISSIRODGKTFIDFKYNDKLPYISNPNYKVNYYAVTKENTIIINSENGDISTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735
RESULT 8
US-09-848-909-7
; Sequence 7, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 06742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-7
Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIVKKSDEVTATSDNHNVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIVKKSDEVTATSDNHNVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQJPELKOKSSNSRKKRSTASGTPVDPDRN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQJPELKOKSSNSRKKRSTASGTPVDPDRN 180
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASPFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASPFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYNNVLPITSLVGNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYNNVLPITSLVGNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTITWNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKDASSPTITWNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKAFGNEPENG 540
QY 541 QYQKDI TEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAKONILIRDRKPHYDR 600

```
Db 541 QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKONLILIRKCFHYDR 600
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660
Qy 661 DMLNSSLRODGTFFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNSSLRODGTFFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Qy 721 IKKILFSKKGYEIG 735
Db 721 IKKILFSKKGYEIG 735

RESULT 9
US-09-848-909-8
; Sequence 8, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-8

Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVITSSDNLQPLPELKQSSNKRKRSTAGPTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVITSSDNLQPLPELKQSSNKRKRSTAGPTVPDRN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 421 LNAQKDASSPTITMNNYQFLEKTKQLRDTDOVYGNATYTFENGVRVDTGNSWSEV 480
Db 421 LNAQKDASSPTITMNNYQFLEKTKQLRDTDOVYGNATYTFENGVRVDTGNSWSEV 480
```

```
Qy 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGL 540
Db 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGL 540
Qy 541 QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKONLILIRKCFHYDR 600
Db 541 QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKONLILIRKCFHYDR 600
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660
Qy 661 DMLNSSLRODGTFFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNSSLRODGTFFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Qy 721 IKKILFSKKGYEIG 735
Db 721 IKKILFSKKGYEIG 735

RESULT 10
US-09-848-909-9
; Sequence 9, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-9

Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVITSSDNLQPLPELKQSSNKRKRSTAGPTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVITSSDNLQPLPELKQSSNKRKRSTAGPTVPDRN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYNTGTAPIYVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
```

```

Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKQDASSTPTITMNYNQFLELEKTKQLRLDQVYGNIAATYVNGRVRVDTGSNWSEV 480
Db 421 LNAQDFFSSTPTITMNYNQFLELEKTKQLRLDQVYGNIAATYVNGRVRVDTGSNWSEV 480
Qy 481 LPOIQETARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAFGFNEPENGNL 540
Db 481 LPOIQETARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAFGFNEPENGNL 540
Qy 541 QYQGKDIITEFDNFDOQTSONIKNQLAELNATNITVLDKIKLAKMKNILIRDKRPHYDR 600
Db 541 QYQGKDIITEFDNFDOQTSONIKNQLAELNATNITVLDKIKLAKMKNILIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGTTFDFKKNYDKLPLYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGTTFDFKKNYDKLPLYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

```

RESULT 11

```

US-09-848-909-10
; Sequence 10, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-10

```

```

Query Match          99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLQJPELKQKSSNSRKKRSTAGPTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLQJPELKQKSSNSRKKRSTAGPTVPDRDN 180
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPYSDPEKVT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPYSDPEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSQNTDSETRTISKNTSRTHT 300

```

```

Qy 301 SEVHGNAEVHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKQDASSTPTITMNYNQFLELEKTKQLRLDQVYGNIAATYVNGRVRVDTGSNWSEV 480
Db 421 LNAQDFFSSTPTITMNYNQFLELEKTKQLRLDQVYGNIAATYVNGRVRVDTGSNWSEV 480
Qy 481 LPOIQETARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAFGFNEPENGNL 540
Db 481 LPOIQETARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAFGFNEPENGNL 540
Qy 541 QYQGKDIITEFDNFDOQTSONIKNQLAELNATNITVLDKIKLAKMKNILIRDKRPHYDR 600
Db 541 QYQGKDIITEFDNFDOQTSONIKNQLAELNATNITVLDKIKLAKMKNILIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGTTFDFKKNYDKLPLYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGTTFDFKKNYDKLPLYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

```

RESULT 12

```

US-09-848-909-11
; Sequence 11, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-11

```

```

Query Match          99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLQJPELKQKSSNSRKKRSTAGPTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLQJPELKQKSSNSRKKRSTAGPTVPDRDN 180
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPYSDPEKVT 240

```

Db 181 DGIPDSLEVEGYTVDVKNKFTFLSPWISNHEKGLTKYKSPKSTASDPYSDFEKT 240
Qy 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEQSTQNTDSETRTISKNTSRHT 300
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEQSTQNTDSETRTISKNTSRHT 300
Qy 301 SEVHGNAEVAHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAETMGLNTADTARL 360
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENOLSQLAPNNYPSKQLAPIA 420
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENOLSQLAPNNYPSKQLAPIA 420
Qy 421 LNAQKDASSPTITWYNOFLFKTQRLDTPQVYGNATYNFENGRVRVDTGSNWSEV 480
Db 421 LNAQKDASSPTITWYNOFLFKTQRLDTPQVYGNATYNFENGRVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWTLKEALKIAGFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWTLKEALKIAGFNEPNGNL 540
Qy 541 QYQKDIETEDFNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIETEDFNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
Qy 601 NNAIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIYEIEDTEGLKEVINDRY 660
Db 601 NNAIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIYEIEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRODGTFFIDFKYNDKPLIYISNPKNVYAVTKNTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGTFFIDFKYNDKPLIYISNPKNVYAVTKNTIINPSENGDTSTNG 720
Qy 721 IKKILIPSKKGYEIG 735
Db 721 IKKILIPSKKGYEIG 735

RESULT 13
US-09-848-909-12
; Sequence 12, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; TYPE: PRT
; LENGTH: 736
; ORGANISM: Bacillus anthracis
US-09-848-909-12

Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYGTFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYGTFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKSDSEYTFATSDNHNVTMWYDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATSDNHNVTMWYDDQEVINKASNNKIRLEKGRLYQIKIY 120

Qy 121 QRENPTKGLDFFKLYWTDSONKKEVIVSSDNLQLPKQKSSNSRKKRSTASGFTVDPDRN 180
Db 121 QRENPTKGLDFFKLYWTDSONKKEVIVSSDNLQLPKQKSSNSRKKRSTASGFTVDPDRN 180
Qy 181 DGIPDSLEVEGYTVDVKNKFTFLSPWISNHEKGLTKYKSPKSTASDPYSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKFTFLSPWISNHEKGLTKYKSPKSTASDPYSDFEKT 240
Qy 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEQSTQNTDSETRTISKNTSRHT 300
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEQSTQNTDSETRTISKNTSRHT 300
Qy 301 SEVHGNAEVAHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAETMGLNTADTARL 360
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENOLSQLAPNNYPSKQLAPIA 420
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENOLSQLAPNNYPSKQLAPIA 420
Qy 421 LNAQKDASSPTITWYNOFLFKTQRLDTPQVYGNATYNFENGRVRVDTGSNWSEV 480
Db 421 LNAQKDASSPTITWYNOFLFKTQRLDTPQVYGNATYNFENGRVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWTLKEALKIAGFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWTLKEALKIAGFNEPNGNL 540
Qy 541 QYQKDIETEDFNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIETEDFNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
Qy 601 NNAIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIYEIEDTEGLKEVINDRY 660
Db 601 NNAIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIYEIEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRODGTFFIDFKYNDKPLIYISNPKNVYAVTKNTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGTFFIDFKYNDKPLIYISNPKNVYAVTKNTIINPSENGDTSTNG 720
Qy 721 IKKILIPSKKGYEIG 735
Db 721 IKKILIPSKKGYEIG 735

RESULT 14
US-09-848-909-13
; Sequence 13, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; TYPE: PRT
; LENGTH: 736
; ORGANISM: Bacillus anthracis
US-09-848-909-13

Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYGTFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60

```
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGLSIPSSSELENIPSENGYF 60
Qy 61 QSAIWSGFTKVKKSDDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFTKVKKSDDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPELKQKSSNSRKRSTASDPSYFVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPELKQKSSNSRKRSTASDPSYFVDRDN 180
Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKMWSTASDPSYFVDRDN 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKMWSTASDPSYFVDRDN 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDSASPTIMTNVNOFLELEKTQKRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKDSASPTIMTNVNOFLELEKTQKRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540
Qy 541 QYQKGDITEFDNFDDQSTQNIKNQLAELNATNIYTVLDKIKNAKMNIILIRDKRPHYDR 600
Db 541 QYQKGDITEFDNFDDQSTQNIKNQLAELNATNIYTVLDKIKNAKMNIILIRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
Qy 661 DMLNISLRODQKTFIDFKYNDKLPYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISLRODQKTFIDFKYNDKLPYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

RESULT 15

```
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication NC. US20020039586A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/050002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRN
; ORGANISM: Bacillus anthracis
US-09-848-909-14
```

Search completed: May 3, 2004, 20:02:13
Job time : 38.6393 secs

```
Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2,9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGLSIPSSSELENIPSENGYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGLSIPSSSELENIPSENGYF 60
Qy 61 QSAIWSGFTKVKKSDDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFTKVKKSDDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPELKQKSSNSRKRSTASDPSYFVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPELKQKSSNSRKRSTASDPSYFVDRDN 180
Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKMWSTASDPSYFVDRDN 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKMWSTASDPSYFVDRDN 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDSASPTIMTNVNOFLELEKTQKRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKDSASPTIMTNVNOFLELEKTQKRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540
Qy 541 QYQKGDITEFDNFDDQSTQNIKNQLAELNATNIYTVLDKIKNAKMNIILIRDKRPHYDR 600
Db 541 QYQKGDITEFDNFDDQSTQNIKNQLAELNATNIYTVLDKIKNAKMNIILIRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
Qy 661 DMLNISLRODQKTFIDFKYNDKLPYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISLRODQKTFIDFKYNDKLPYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

1	3772	100.0	735	1	PCT-US03-35733-13	Sequence 13, Appl
2	3772	100.0	735	23	US-09-848-909A-13	Sequence 13, Appl
3	3766	99.8	735	1	PCT-US03-35733-10	Sequence 10, Appl
4	3766	99.8	735	23	US-09-848-909A-10	Sequence 10, Appl
5	3763	99.8	735	1	PCT-US03-19786-4	Sequence 4, Appl
6	3760	99.7	735	1	PCT-US03-35733-2	Sequence 2, Appl
7	3760	99.7	735	23	US-09-848-909A-2	Sequence 2, Appl
8	3759	99.7	735	1	PCT-US03-35733-8	Sequence 8, Appl
9	3759	99.7	735	1	PCT-US03-35733-9	Sequence 9, Appl
10	3759	99.7	735	23	US-09-848-909A-8	Sequence 9, Appl
11	3759	99.7	735	23	US-09-848-909A-9	Sequence 9, Appl
12	3755	99.5	735	1	PCT-US03-35733-7	Sequence 7, Appl
13	3755	99.5	735	1	PCT-US03-35733-23	Sequence 23, Appl
14	3755	99.5	735	23	US-09-848-909A-7	Sequence 7, Appl
15	3755	99.5	735	23	US-09-848-909A-23	Sequence 23, Appl
16	3754	99.5	735	1	PCT-US03-35733-4	Sequence 4, Appl
17	3754	99.5	735	1	PCT-US03-35733-6	Sequence 6, Appl
18	3754	99.5	735	1	PCT-US03-35733-18	Sequence 18, Appl
19	3754	99.5	735	23	US-09-848-909A-4	Sequence 4, Appl
20	3754	99.5	735	23	US-09-848-909A-6	Sequence 6, Appl
21	3754	99.5	735	23	US-09-848-909A-18	Sequence 18, Appl
22	3753	99.5	735	1	PCT-US03-35733-5	Sequence 5, Appl
23	3753	99.5	735	1	PCT-US03-35733-11	Sequence 11, Appl
24	3753	99.5	735	1	PCT-US03-35733-17	Sequence 17, Appl
25	3753	99.5	735	1	PCT-US03-35733-19	Sequence 19, Appl
26	3753	99.5	735	1	PCT-US03-35733-20	Sequence 20, Appl
27	3753	99.5	735	1	PCT-US03-35733-21	Sequence 21, Appl
28	3753	99.5	735	22	US-09-791-537-43735	Sequence 43735, A
29	3753	99.5	735	23	US-09-848-909A-5	Sequence 5, Appl
30	3753	99.5	735	23	US-09-848-909A-11	Sequence 11, Appl
31	3753	99.5	735	23	US-09-848-909A-17	Sequence 17, Appl
32	3753	99.5	735	23	US-09-848-909A-19	Sequence 19, Appl
33	3753	99.5	735	23	US-09-848-909A-20	Sequence 20, Appl
34	3753	99.5	735	23	US-09-848-909A-21	Sequence 21, Appl
35	3753	99.5	735	30	US-10-410-647-30	Sequence 30, Appl
36	3753	99.5	736	1	PCT-US01-14372A-1	Sequence 1, Appl
37	3753	99.5	736	1	PCT-US01-14372A-2	Sequence 2, Appl
38	3753	99.5	736	1	PCT-US01-14372A-3	Sequence 3, Appl
39	3753	99.5	736	1	PCT-US01-14372A-4	Sequence 4, Appl
40	3753	99.5	736	1	PCT-US01-14372A-5	Sequence 5, Appl
41	3753	99.5	736	1	PCT-US01-14372A-6	Sequence 6, Appl
42	3753	99.5	736	1	PCT-US01-14372A-7	Sequence 7, Appl
43	3753	99.5	736	1	PCT-US01-14372A-8	Sequence 8, Appl
44	3753	99.5	736	1	PCT-US01-14372A-9	Sequence 9, Appl
45	3753	99.5	736	1	PCT-US01-14372A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
PCT-US03-35733-13
; Sequence 13, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-13

Query Match 100.0%; Score 3772; DB 1; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.8e-298;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 175.546 Seconds
(without alignments)
4086.665 Million cell updates/sec

Title: US-09-848-909A-13
Perfect score: 3772
Sequence: 1 EVKQENRLNSESSESQGLL.....TSTNGIKILLIFSKGYBIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues
Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_Main:*

- 1: /cgn2_6/prodata/2/paa/US087_COMB.pcp.*
- 2: /cgn2_6/prodata/2/paa/US06_COMB.pcp.*
- 3: /cgn2_6/prodata/2/paa/US07_COMB.pcp.*
- 4: /cgn2_6/prodata/2/paa/US087_COMB.pcp.*
- 5: /cgn2_6/prodata/2/paa/US081_COMB.pcp.*
- 6: /cgn2_6/prodata/2/paa/US082_COMB.pcp.*
- 7: /cgn2_6/prodata/2/paa/US083_COMB.pcp.*
- 8: /cgn2_6/prodata/2/paa/US084_COMB.pcp.*
- 9: /cgn2_6/prodata/2/paa/US085_COMB.pcp.*
- 10: /cgn2_6/prodata/2/paa/US086_COMB.pcp.*
- 11: /cgn2_6/prodata/2/paa/US087_COMB.pcp.*
- 12: /cgn2_6/prodata/2/paa/US088_COMB.pcp.*
- 13: /cgn2_6/prodata/2/paa/US089_COMB.pcp.*
- 14: /cgn2_6/prodata/2/paa/US090_COMB.pcp.*
- 15: /cgn2_6/prodata/2/paa/US091_COMB.pcp.*
- 16: /cgn2_6/prodata/2/paa/US092_COMB.pcp.*
- 17: /cgn2_6/prodata/2/paa/US093_COMB.pcp.*
- 18: /cgn2_6/prodata/2/paa/US094_COMB.pcp.*
- 19: /cgn2_6/prodata/2/paa/US095_COMB.pcp.*
- 20: /cgn2_6/prodata/2/paa/US096_COMB.pcp.*
- 21: /cgn2_6/prodata/2/paa/US097A_COMB.pcp.*
- 22: /cgn2_6/prodata/2/paa/US097B_COMB.pcp.*
- 23: /cgn2_6/prodata/2/paa/US098_COMB.pcp.*
- 24: /cgn2_6/prodata/2/paa/US099A_COMB.pcp.*
- 25: /cgn2_6/prodata/2/paa/US099B_COMB.pcp.*
- 26: /cgn2_6/prodata/2/paa/US100_COMB.pcp.*
- 27: /cgn2_6/prodata/2/paa/US101_COMB.pcp.*
- 28: /cgn2_6/prodata/2/paa/US102_COMB.pcp.*
- 29: /cgn2_6/prodata/2/paa/US103_COMB.pcp.*
- 30: /cgn2_6/prodata/2/paa/US104_COMB.pcp.*
- 31: /cgn2_6/prodata/2/paa/US106_COMB.pcp.*
- 32: /cgn2_6/prodata/2/paa/US107_COMB.pcp.*
- 33: /cgn2_6/prodata/2/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

QY 1 EVKQENRLNSESSESSQGLLGYFSDLNFAQPMVVTSTTTGDLSPSSSELENIIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFAQPMVVTSTTTGDLSPSSSELENIIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKKRSTASAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKKRSTASAGTVPDRDN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKT 240
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKASSTPIITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGNSWSEV 480
Db 421 LNAQKASSTPIITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIETEDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKONILIRDKRPHYDR 600
Db 541 QYQKDIETEDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKONILIRDKRPHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
QY 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 2

US-09-848-909A-13
; Sequence 13, Application US/09848909A

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909A

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-05-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909A-13

Query Match 100.0%; Score 3772; DB 23; Length 735;

Best Local Similarity 100.0%; Pred. No. 1.8e-238;

Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLGYFSDLNFAQPMVVTSTTTGDLSPSSSELENIIPSENQYF 60

Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFAQPMVVTSTTTGDLSPSSSELENIIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKKRSTASAGTVPDRDN 180

Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKKRSTASAGTVPDRDN 180

QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKT 240

Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKT 240

QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

QY 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Db 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKASSTPIITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGNSWSEV 480

Db 421 LNAQKASSTPIITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGNSWSEV 480

QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

QY 541 QYQKDIETEDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKONILIRDKRPHYDR 600

Db 541 QYQKDIETEDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKONILIRDKRPHYDR 600

QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660

Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660

QY 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

Db 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

QY 721 IKKILIFSCKGYEIG 735

Db 721 IKKILIFSCKGYEIG 735

RESULT 3

PCT-US03-35733-10

; Sequence 10, Application PC/TUS0335733

; GENERAL INFORMATION:

; APPLICANT: President and Fellows of Harvard College et al.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/072003

; CURRENT APPLICATION NUMBER: PCT/US03/35733

; CURRENT FILING DATE: 2003-10-10

; PRIOR APPLICATION NUMBER: US 60/424,987

; PRIOR FILING DATE: 2002-11-08

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0


```
; SEQ ID NO 10
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-10

Query Match          99.8%; Score 3766; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 5.5e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLLYGFFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSESSQGLLYGFFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNLQJPELKQKSSNSRKRSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNLQJPELKQKSSNSRKRSTAGTVPDRDN 180
QY 181 DGI PDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
DB 181 DGI PDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSSTPTITWNYNQFLEKTKQLRLDTPQVGNATYNFENGVRVDTGSNWSEV 480
DB 421 LNAQKDSSTPTITWNYNQFLEKTKQLRLDTPQVGNATYNFENGVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKOLNLVERRIA AVNPSDPLETTKPDMTLKEALKAFGFNEPENG 540
DB 481 LPOIQTETARIIFNGKOLNLVERRIA AVNPSDPLETTKPDMTLKEALKAFGFNEPENG 540
QY 541 QYQKDI TEFDNFDQQTSONIKNQLAELNATNIYTVLDKIKLNAKKNILIRDKRFHYDR 600
DB 541 QYQKDI TEFDNFDQQTSONIKNQLAELNATNIYTVLDKIKLNAKKNILIRDKRFHYDR 600
QY 601 NNIAVGADSVVKEAHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADSVVKEAHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 4
US-09-848-909A-10
; Sequence 10, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; OF BACTERIAL INFECTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
```

```
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-10

Query Match          99.8%; Score 3766; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 5.5e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLLYGFFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSESSQGLLYGFFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNLQJPELKQKSSNSRKRSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNLQJPELKQKSSNSRKRSTAGTVPDRDN 180
QY 181 DGI PDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
DB 181 DGI PDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSSTPTITWNYNQFLEKTKQLRLDTPQVGNATYNFENGVRVDTGSNWSEV 480
DB 421 LNAQKDSSTPTITWNYNQFLEKTKQLRLDTPQVGNATYNFENGVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKOLNLVERRIA AVNPSDPLETTKPDMTLKEALKAFGFNEPENG 540
DB 481 LPOIQTETARIIFNGKOLNLVERRIA AVNPSDPLETTKPDMTLKEALKAFGFNEPENG 540
QY 541 QYQKDI TEFDNFDQQTSONIKNQLAELNATNIYTVLDKIKLNAKKNILIRDKRFHYDR 600
DB 541 QYQKDI TEFDNFDQQTSONIKNQLAELNATNIYTVLDKIKLNAKKNILIRDKRFHYDR 600
QY 601 NNIAVGADSVVKEAHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADSVVKEAHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 5
PCT-US03-19786-4
; Sequence 4, Application PC/TUS0319786
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
```


; TITLE OF INVENTION: Modified Shine Dalgarno Sequences and Methods of Use Thereof
; FILE REFERENCE: PV59SPCT
; CURRENT APPLICATION NUMBER: PCT/US03/19786
; CURRENT FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Mature PA sequence including an E7B signal sequence
PCT-US03-19786-4

Query Match 99.8%; Score 3763; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 9.7e-298;
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy	1	EVKQENRLNESSSQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Db	1	EVKQENRLNESSSQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Qy	61	QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
Db	61	QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
Qy	121	QRENPEKGLDFKLYWTDSONKKEVSSDNLQIPELKQKSSNRKKGSTAGPTVPDRDN	180
Db	121	QRENPEKGLDFKLYWTDSONKKEVSSDNLQIPELKQKSSNRKKGSTAGPTVPDRDN	180
Qy	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPVSDPEKVT	240
Db	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPVSDPEKVT	240
Qy	241	GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETRTISKNTSTSRTH	300
Db	241	GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETRTISKNTSTSRTH	300
Qy	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Qy	361	NANRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA	420
Db	361	NANRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA	420
Qy	421	LNAQKDSSTPIITWYNOQFLEKTKQLRLDQVYGNIAITYNFENGVRVDTGSNWSEV	480
Db	421	LNAQKDSSTPIITWYNOQFLEKTKQLRLDQVYGNIAITYNFENGVRVDTGSNWSEV	480
Qy	481	LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG	540
Db	481	LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG	540
Qy	541	QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR	600
Db	541	QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR	600
Qy	601	NNIYAVGADESIVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVIN	660
Db	601	NNIYAVGADESIVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVIN	660
Qy	661	DMLNISSLRQDKTFIDFKYNDKPLIYISNPYKNVYAVTKNTIINPSENGDTSTNG	720
Db	661	DMLNISSLRQDKTFIDFKYNDKPLIYISNPYKNVYAVTKNTIINPSENGDTSTNG	720
Qy	721	IKKILIFSKKGYEIG 735	
Db	721	IKKILIFSKKGYEIG 735	

RESULT 6
PCT-US03-35733-2

; Sequence 2, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-2

Query Match 99.7%; Score 3760; DB 1; Length 735; Best Local Similarity 99.7%; Pred. No. 1.7e-297; Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Qy	1	EVKQENRLNESSSQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
	Db	1	EVKQENRLNESSSQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
	Qy	61	QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
	Db	61	QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
	Qy	121	QRENPEKGLDFKLYWTDSONKKEVSSDNLQIPELKQKSSNRKKGSTAGPTVPDRDN	180
	Db	121	QRENPEKGLDFKLYWTDSONKKEVSSDNLQIPELKQKSSNRKKGSTAGPTVPDRDN	180
	Qy	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPVSDPEKVT	240
	Db	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPVSDPEKVT	240
	Qy	241	GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETRTISKNTSTSRTH	300
	Db	241	GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETRTISKNTSTSRTH	300
	Qy	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
	Db	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
	Qy	361	NANRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA	420
	Db	361	NANRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA	420
	Qy	421	LNAQKDSSTPIITWYNOQFLEKTKQLRLDQVYGNIAITYNFENGVRVDTGSNWSEV	480
	Db	421	LNAQKDSSTPIITWYNOQFLEKTKQLRLDQVYGNIAITYNFENGVRVDTGSNWSEV	480
	Qy	481	LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG	540
	Db	481	LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG	540
	Qy	541	QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR	600
	Db	541	QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR	600
	Qy	601	NNIYAVGADESIVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVIN	660
	Db	601	NNIYAVGADESIVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVIN	660
	Qy	661	DMLNISSLRQDKTFIDFKYNDKPLIYISNPYKNVYAVTKNTIINPSENGDTSTNG	720
	Db	661	DMLNISSLRQDKTFIDFKYNDKPLIYISNPYKNVYAVTKNTIINPSENGDTSTNG	720
	Qy	721	IKKILIFSKKGYEIG 735	
	Db	721	IKKILIFSKKGYEIG 735	

```
RESULT 7
US-09-848-909A-2
; Sequence 2, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,900
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-2

Query Match          99.7%; Score 3760; DB 23; Length 735;
Best Local Similarity 99.7%; Pred. No. 1.7e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKRSSTAGTVPDRDN 180
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
DB 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSOILAPNNYPSKNLAPIA 420
DB 361 NANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSOILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSSTPTITWYVNFQLEKTKQLRLDQVYGNATYFNGRVRVDTGNSNNEV 480
DB 421 LNAQKDSSTPTITWYVNFQLEKTKQLRLDQVYGNATYFNGRVRVDTGNSNNEV 480
QY 481 LPQOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENGNL 540
DB 481 LPQOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENGNL 540
QY 541 QYQKDIITEFDNFDQQTSONIKQLAELNATNIYTVLDKIKLNAKKNILIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDQQTSONIKQLAELNATNIYTVLDKIKLNAKKNILIRDKRPHYDR 600
QY 601 NNIAVGADESIVKKAHREVINSTEGLLLNIDKDIRKILSGYVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESIVKKAHREVINSTEGLLLNIDKDIRKILSGYVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKFLYISNPNYKVNVAVTKEKTIINPSENGDSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKYNDKFLYISNPNYKVNVAVTKEKTIINPSENGDSTNG 720

RESULT 8
PCT-US03-35733-8
; Sequence 8, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-8

Query Match          99.7%; Score 3759; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 2.1e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKRSSTAGTVPDRDN 180
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
DB 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSOILAPNNYPSKNLAPIA 420
DB 361 NANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSOILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSSTPTITWYVNFQLEKTKQLRLDQVYGNATYFNGRVRVDTGNSNNEV 480
DB 421 LNAQKDSSTPTITWYVNFQLEKTKQLRLDQVYGNATYFNGRVRVDTGNSNNEV 480
QY 481 LPQOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENGNL 540
DB 481 LPQOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENGNL 540
QY 541 QYQKDIITEFDNFDQQTSONIKQLAELNATNIYTVLDKIKLNAKKNILIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDQQTSONIKQLAELNATNIYTVLDKIKLNAKKNILIRDKRPHYDR 600
QY 601 NNIAVGADESIVKKAHREVINSTEGLLLNIDKDIRKILSGYVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESIVKKAHREVINSTEGLLLNIDKDIRKILSGYVEIEDTEGLKEVINDRY 660
```

Db 601 NNIAGADESVKAEHREVINSTEGLLLNIDKIRKILSGYVIEIEDTEGLKEVINDRY 660
Qy 661 DMLNLSLRODQGTFFIDFKKYNNDKLPYISNPNYKNNVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNLSLRODQGTFFIDFKKYNNDKLPYISNPNYKNNVAVTKENTIIINPSNGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735
RESULT 9
PCT-US03-35733-9
; Sequence 9, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-9
Query Match 99.7%; Score 3759; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 2.1e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSGDEYTFATSDAHVVTMWDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSGDEYTFATSDAHVVTMWDDQEVINKASNNKIRLEKGRLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPPELKQKSNRKKRSTSGAGTPVDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPPELKQKSNRKKRSTSGAGTPVDRDN 180
Qy 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENQSLQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDDASSTPTIMYNOFLEKTKQLRLDQDVYGNATYNFENGVRVDTGSGNWSEV 480
Db 421 LNAQKDDASSTPTIMYNOFLEKTKQLRLDQDVYGNATYNFENGVRVDTGSGNWSEV 480
Qy 481 LPOIETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
Db 481 LPOIETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
Qy 541 QYQKIDITEFFDNFQOOTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600

Db 541 QYQKIDITEFFDNFQOOTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
Qy 601 NNIAGADESVKAEHREVINSTEGLLLNIDKIRKILSGYVIEIEDTEGLKEVINDRY 660
Db 601 NNIAGADESVKAEHREVINSTEGLLLNIDKIRKILSGYVIEIEDTEGLKEVINDRY 660
Qy 661 DMLNLSLRODQGTFFIDFKKYNNDKLPYISNPNYKNNVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNLSLRODQGTFFIDFKKYNNDKLPYISNPNYKNNVAVTKENTIIINPSNGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735
RESULT 10
US-09-848-909A-8
; Sequence 8, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-8
Query Match 99.7%; Score 3759; DB 23; Length 735;
Best Local Similarity 99.7%; Pred. No. 2.1e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSGDEYTFATSDAHVVTMWDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSGDEYTFATSDAHVVTMWDDQEVINKASNNKIRLEKGRLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPPELKQKSNRKKRSTSGAGTPVDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPPELKQKSNRKKRSTSGAGTPVDRDN 180
Qy 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENQSLQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDDASSTPTIMYNOFLEKTKQLRLDQDVYGNATYNFENGVRVDTGSGNWSEV 480
Db 421 LNAQKDDASSTPTIMYNOFLEKTKQLRLDQDVYGNATYNFENGVRVDTGSGNWSEV 480
Qy 481 LPOIETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540

Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPGL 540
QY 541 QYQKDIITEFDNFQDOTSONIKQLAELNATNIYVLDKIKLNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFQDOTSONIKQLAELNATNIYVLDKIKLNAKNNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 11
US-09-848-909A-9
; Sequence 9, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-9

Query Match 99.7%; Score 3759; DB 23; Length 735;
Best Local Similarity 99.7%; Pred. No. 2, 1e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYGYFSDLNQFAPMVVTSSTTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLLYGYFSDLNQFAPMVVTSSTTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYEYTPATSDADNHVTMWVDDQEVINKASNNKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYEYTPATSDADNHVTMWVDDQEVINKASNNKIRLEKRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNRKCRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNRKCRSTASGTPVDRDN 180
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEYV 240
Db 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEYV 240
QY 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVNTGTAPIYVNLPTTSVLGKNTLATIRAKENQLSQILAPNNYPSKMLAPIA 420
Db 361 NANIRVNTGTAPIYVNLPTTSVLGKNTLATIRAKENQLSQILAPNNYPSKMLAPIA 420

QY 421 LNAQKASSTPIITMYNQFLEBKTQKLRDLDQVYGNATYNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKASSTPIITMYNQFLEBKTQKLRDLDQVYGNATYNFENGRVVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPGL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPGL 540
QY 541 QYQKDIITEFDNFQDOTSONIKQLAELNATNIYVLDKIKLNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFQDOTSONIKQLAELNATNIYVLDKIKLNAKNNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 12
PCT-US03-35733-7
; Sequence 7, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-7

Query Match 99.5%; Score 3755; DB 1; Length 735;
Best Local Similarity 99.6%; Pred. No. 4, 4e-297;
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYGYFSDLNQFAPMVVTSSTTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLLYGYFSDLNQFAPMVVTSSTTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYEYTPATSDADNHVTMWVDDQEVINKASNNKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYEYTPATSDADNHVTMWVDDQEVINKASNNKIRLEKRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNRKCRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNRKCRSTASGTPVDRDN 180
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEYV 240
Db 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEYV 240
QY 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NAMIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NAMIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKADASSPTITMNNYQFLELEKTKQRLDQVYGNATYNFENGRVVDVDTGSNWSEV 480
DB 421 LNAQEDSSPTITMNNYQFLELEKTKQRLDQVYGNATYNFENGRVVDVDTGSNWSEV 480
QY 481 LPOIQETTARIIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENGNL 540
DB 481 LPOIQETTARIIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENGNL 540
QY 541 QYQKDIITEPDRFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
DB 541 QYQKDIITEPDRFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
DB 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISLRODQGTFFDFKKNYDKLPLYISNPNYKVNVAVTKENTIINPSENGTSTNG 720
DB 661 DMLNISLRODQGTFFDFKKNYDKLPLYISNPNYKVNVAVTKENTIINPSENGTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735
RESULT 13
PCT-US03-35733-23
; Sequence 23, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; NAME/KEY: VARIANT
; LOCATION: 427
; OTHER INFORMATION: Xaa = any amino acid except Phe
PCT-US03-35733-23
Query Match 99.5%; Score 3755; DB 1; Length 735;
Best Local Similarity 99.6%; Pred. No. 4.4e-297;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENYF 60
DB 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNNKIRLEKRLYQIKQY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNNKIRLEKRLYQIKQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQJPELKQKSSNRKKRSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQJPELKQKSSNRKKRSTAGTVPDRDN 180
QY 181 DGIPTDSLEVEGYTVDVKNKRTFUSPWISNHEKKGITKYSPEKWKSTASDPSPEKVT 240
DB 181 DGIPTDSLEVEGYTVDVKNKRTFUSPWISNHEKKGITKYSPEKWKSTASDPSPEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEYHASFDDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
DB 301 SEVHGNAEYHASFDDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
QY 361 NAMIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NAMIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKADASSPTITMNNYQFLELEKTKQRLDQVYGNATYNFENGRVVDVDTGSNWSEV 480
DB 421 LNAQDDSSPTITMNNYQFLELEKTKQRLDQVYGNATYNFENGRVVDVDTGSNWSEV 480
QY 481 LPOIQETTARIIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENGNL 540
DB 481 LPOIQETTARIIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENGNL 540
QY 541 QYQKDIITEPDRFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
DB 541 QYQKDIITEPDRFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
DB 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISLRODQGTFFDFKKNYDKLPLYISNPNYKVNVAVTKENTIINPSENGTSTNG 720
DB 661 DMLNISLRODQGTFFDFKKNYDKLPLYISNPNYKVNVAVTKENTIINPSENGTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735
RESULT 14
US-09-848-909A-7
; Sequence 7, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-7
Query Match 99.5%; Score 3755; DB 23; Length 735;
Best Local Similarity 99.6%; Pred. No. 4.4e-297;
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENYF 60
DB 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNNKIRLEKRLYQIKQY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNNKIRLEKRLYQIKQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQJPELKQKSSNRKKRSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQJPELKQKSSNRKKRSTAGTVPDRDN 180

181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKVT 240
181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKVT 240
241 GRIDKNVSPPEARHPVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
241 GRIDKNVSPPEARHPVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPPSKNLAPIA 420
361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPPSKNLAPIA 420
421 LNAQKDSSTPTIMNYNQFLFKYQLRDLTDQVYGNATYNFENGVRVDTGNSWSEV 480
421 LNAQKDSSTPTIMNYNQFLFKYQLRDLTDQVYGNATYNFENGVRVDTGNSWSEV 480
481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
541 QYQKDIETFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 600
541 QYQKDIETFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 600
601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720
661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720
721 IKKILIFSKKGYEIG 735
721 IKKILIFSKKGYEIG 735

RESULT 15
US-09-848-909A-23
; Sequence 23 Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 427
; OTHER INFORMATION: Xaa = any amino acid except Phe
US-09-848-909A-23

Query Match 99.5%; Score 3755; DB 23; Length 735;
Best Local Similarity 99.6%; Pred. No. 4.4e-297;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 EVKQENRLLNSESQGLLYFFDLNFQAPMVVTSSTTGLDLSIPSELENIPSENQYF 60

1 EVKQENRLLNSESQGLLYFFDLNFQAPMVVTSSTTGLDLSIPSELENIPSENQYF 60
61 QSAINSGFTKVKKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGLYQIKIY 120
61 QSAINSGFTKVKKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGLYQIKIY 120
121 QRENPTKGLDFKLYWTDSONKKEVSSONLQPELKQKSSNSRKRKSTASAGTVPDRDN 180
121 QRENPTKGLDFKLYWTDSONKKEVSSONLQPELKQKSSNSRKRKSTASAGTVPDRDN 180
181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKVT 240
181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKVT 240
241 GRIDKNVSPPEARHPVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
241 GRIDKNVSPPEARHPVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPPSKNLAPIA 420
361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPPSKNLAPIA 420
421 LNAQKDSSTPTIMNYNQFLFKYQLRDLTDQVYGNATYNFENGVRVDTGNSWSEV 480
421 LNAQKDSSTPTIMNYNQFLFKYQLRDLTDQVYGNATYNFENGVRVDTGNSWSEV 480
481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
541 QYQKDIETFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 600
541 QYQKDIETFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 600
601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720
661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720
721 IKKILIFSKKGYEIG 735
721 IKKILIFSKKGYEIG 735

Search completed: May 3, 2004, 19:57:44
Job time : 177.546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 13.0855 Seconds
(without alignments)
5403.004 Million cell updates/sec

Title: US-09-848-909A-13
Perfect score: 3772
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKLIFSKGYEIG 735
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:**
1: pir1:**
2: pir2:**
3: pir3:**
4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3753	99.5	764	2 I39934	protective antigen
2	868	23.0	875	2 I40862	iota toxin compone
3	235.5	6.2	192	2 I39933	cryptic protein -
4	235.5	6.2	204	2 G59104	hypothetical prote
5	202	5.4	4688	2 F82885	hypothetical prote
6	190.5	5.1	2401	2 T28676	hypothetical prote
7	189.5	5.0	2523	2 B64635	toxin-like outer m
8	186.5	4.9	1125	2 E90598	membrane nuclease,
9	184.5	4.9	1639	2 S05603	major merozoite su
10	183	4.9	6713	2 B89921	hypothetical prote
11	181	4.8	1302	1 JC6009	surface-located me
12	181	4.8	2269	2 T28677	rhoetry protein -
13	179.5	4.8	1072	2 A86827	hypothetical prote
14	179	4.7	1193	2 S68218	botulinum neurotox
15	179	4.7	1227	2 C97033	uncharacterized pr
16	178	4.7	1635	2 A10452	hemolysin [importe
17	177.5	4.7	2399	2 H71879	toxin-like outer m
18	177	4.7	4919	2 T31105	hypothetical prote
19	176	4.7	752	2 G90599	hypothetical prote
20	174.5	4.6	1385	2 T30822	lmp1 protein - Myc
21	173	4.6	1939	2 T18372	repeat organellar
22	173	4.6	4152	2 T31102	filamentous hemag
23	172	4.6	1837	2 T41023	probable nuclear p
24	170	4.5	1308	2 E71622	probable membrane
25	169.5	4.5	2340	2 B71704	cell surface antig
26	169	4.5	1631	1 SAZOK1	major merozoite su
27	169	4.5	1658	2 S55101	hypothetical prote
28	169	4.5	5005	2 F82884	hypothetical prote
29	168.5	4.5	821	2 S67087	hypothetical prote

ALIGNMENTS

RESULT 1

I39934
protective antigen precursor - Bacillus anthracis plasmid
C:Species: Bacillus anthracis
C:Date: 19-Jul-1996 #sequence, revision 19-Jul-1996 #text_change 01-Dec-2000
C:Accession: I39934; S69160; F59104
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A>Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr
A:Reference number: I39933; MUID:89172073; PMID:3148491
A:Accession: I39934
A>Status: preliminary; translated from GE/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-764 <RES>
A:Cross-references: GB:M22589; NID:G143280; PIDN:AAA2637.1; PID:G143282
R:Friedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klimpel, K.R.; Birch, N.P.; Loh, Y.P.
Arch. Biochem. Biophys. 316, 5-13, 1995
A>Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SE
A:Reference number: S69160; MUID:95142670; PMID:7840657
A:Accession: S69160
A:Molecule type: protein
A:Residues: 197-202 <PRI>
R:Okimura, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle
J. Bacteriol. 181, 8509-8515, 1999
A>Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbor
A:Reference number: A59091; MUID:99445483; PMID:10515943
A:Accession: F59104
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-313, 'Q', 315-764 <OKI>
A:Cross-references: GB:AF065404; NID:G4894216; PIDN:AAD32414.1; PID:G4894326
A:Experimental source: strain Sterne
A>Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid
C:Genetics:
A:Gene: pXOI-110
A:Genome: plasmid
C:Function:
A:Description: three component exotoxin; protective antigen binds to receptors on the s
y active components edema factor or lethal factor; the complex is internalized by recep
C:Keywords: exotoxin
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-196/Domain: propeptide #status predicted <PRO>
F:197-202/Product: protective antigen #status experimental <MAT>

Query Match 99.5%; Score 3753; DB 2; Length 764;
Best Local Similarity 99.6%; Pred. No. 1.5e-182;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 EVKQENRLNSESQGLLYGYPFDLNFQAPMVVTSITGDLSPSELENIPSENOYF 60
|||||
DB 30 EVKQENRLNSESQGLLYGYPFDLNFQAPMVVTSITGDLSPSELENIPSENOYF 89

probable coiled-co
alpha-toxin - Clos
surface membrane p
trse-like protein
TyB protein - yeas
botulinum neurotox
hypothetical prote
TyB protein - yeas
TyB protein - yeas
hypothetical prote
serine proteinase
hypothetical coile
p115 protein - Myc
hypothetical prote
hypothetical prote
actin-interacting

30 168.5 4.5 1115 2 T41342
31 168 4.5 2178 2 S55805
32 167.5 4.4 624 2 PC6003
33 167.5 4.4 853 2 G90559
34 167.5 4.4 1802 2 S52611
35 166 4.4 1276 2 S11455
36 165.5 4.4 786 2 T18469
37 165 4.4 1465 2 S31262
38 165 4.4 1803 2 S56894
39 165 4.4 3724 2 T18427
40 164 4.3 769 2 F89870
41 164 4.3 1957 2 T38077
42 163.5 4.3 979 2 JQ0894
43 163 4.3 1272 2 C90593
44 163 4.3 3216 2 C90538
45 162 4.3 1033 2 T37715


```
QY 61 QSAIWSGFIKVKKSDEYTFATGADNHVTHWVDDQEVINKASNSNKIRLEKGLVQIKIQ 120
DB 90 QSAIWSGFIKVKKSDEYTFATGADNHVTHWVDDQEVINKASNSNKIRLEKGLVQIKIQ 149
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRSTSGPTVPDRDN 180
DB 150 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRSTSGPTVPDRDN 209
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASTASDPDFEKT 240
DB 210 DGPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASTASDPDFEKT 269
QY 241 GRIDKNVSEARHPLVAAPVIVHVDMEIILSKNEDQSTONTSDSTRISKNSTSRHT 300
DB 270 GRIDKNVSEARHPLVAAPVIVHVDMEIILSKNEDQSTONTSDSTRISKNSTSRHT 329
QY 301 SEVHGNAEYHAFDFDGGVSAGFSNSSTVAIDHSLSLAGERTWAEATMGINTADTARL 360
DB 330 SEVHGNAEYHAFDFDGGVSAGFSNSSTVAIDHSLSLAGERTWAEATMGINTADTARL 389
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSOILAPNNYPSKNLAPIA 420
DB 390 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSOILAPNNYPSKNLAPIA 449
QY 421 LNAQKQDASPTPTMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGNSWSEV 480
DB 450 LNAQKQDASPTPTMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGNSWSEV 509
QY 481 LPOIQETTARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAPGFNPNGL 540
DB 510 LPOIQETTARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAPGFNPNGL 569
QY 541 QYQKQDITEPFDNFQDQTSQNIKNQLAELNATNIYVLDKIKNKAKNNILIRDKRFHYDR 600
DB 570 QYQKQDITEPFDNFQDQTSQNIKNQLAELNATNIYVLDKIKNKAKNNILIRDKRFHYDR 629
QY 601 NNIAVGADESVEKAEHREVNSTEGLLNIDKDIRKILSGYVIEIETEGEKEVINDBY 660
DB 630 NNIAVGADESVEKAEHREVNSTEGLLNIDKDIRKILSGYVIEIETEGEKEVINDBY 689
QY 661 DMLNISLRQDGTFFDFKYNKDLPLYISNPNKYVNVAVTKENTIINPSENGTSTNG 720
DB 690 DMLNISLRQDGTFFDFKYNKDLPLYISNPNKYVNVAVTKENTIINPSENGTSTNG 749
QY 721 IKKILIPSKGYEIG 735
DB 750 IKKILIPSKGYEIG 764

RESULT 2
I40862
iota toxin component Ib - Clostridium perfringens
C:Species: Clostridium perfringens
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Oct-1999
C:Accession: I40862; S42774
R:Perelle, S.; Gilbert, M.; Boquet, P.; Popoff, M.R.
Infect Immun 61, 5147-5156, 1993
A:Title: Characterization of Clostridium perfringens iota-toxin genes and expression in
A:Reference number: I40861; MUID:94041637; PMID:8225592
A:Accession: I40862
A>Status: preliminary; translated from GB/EMBL/DDJ
A:Molecule type: DNA
A:Residues: 1-875 <RES>
A:Cross-references: EMBL:X73562; NID:g929031; PIDN:CRA51960.1; PID:g414655

Query Match 23.0%; Score 868; DB 2; Length 875;
Best Local Similarity 31.1%; Pred. No. 2.1e-36;
Matches 252; Conservative 130; Mismatches 279; Indels 150; Gaps 31;

QY 1 EVKQERILLNESSESSOGLGYFSDNLPQAPMVVTSSTGDSIPSSSELENIPSS-NQY 59
DB 35 DTNQKEITNENTLSSNGLMGYFADHFKDLSEMAPIKNGDLKPFEEKVKDLITEDSS 94
```

```
QY 60 FQSAIWSGFIKVKKSDEYTFATGADNHVTHWVDDQEVINKASNSNKIRLEKGLVQIKIQ 119
DB 95 IKIRMTGTIIIPSEDGEYILSDTR-NDVLQINAKGDIK--TLKVNKKGOAYNIRIE 150
QY 120 YQREN-----PTEKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRST 170
DB 151 IQDKNLGSDNLVSP-----KLYW-ELNGKNTVPEENLPEDYSKIDEND----- 195
QY 171 AGTVP-----DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNTH 211
DB 196 --PFIPIPNPFFVFFSAWEDEDLTDNDNIPDAYEKNGYTI---KDSIAVKWDSFA 249
QY 212 EKKGTLKYKSSPEKWSASTASDPDFEKTGRIDKNVSEARHPLVAAPVIVHVDMEIIL 271
DB 250 E-QGYKYVSSYLESTAGDPYTDYKASGSDKA-KLEARDPLVAAPVIVGVMENLII 308
QY 272 SKNEDQSTONTSDSTRISKNSTSRHTSEVHGNAEYHAFDFDGGVSAGFSNSST 331
DB 309 STNEHASS---DQGTVSEATTSKTDANTV-----GVSIAGYQNGFTGN 351
QY 332 VADHS-----LSLAGERTWAEATMGINTADTARLNANIRVYNTGTAPIYVNLPTTSL 383
DB 352 ITTSYHTTNDNSTAVQDSNGESWNTGLSINKGSAYINANVRVYNTGTAPMYKVTPTNL 411
QY 384 VLGNKQTLATIKADENQLSOILAPNNYPSKNLAPIALNAQKQDASPTPTMNYNQFLE 443
DB 412 VL-DGETLATIKAQDNQIGNNLSFNETYPKKGLSPLALATMDQFNARLIPINVDQLKCLD 470
QY 444 KTKQLRLDQVYGNATYNFENGRVVRVDTGNSWSEVLPQIQETTARIIFNGKDLNVER 503
DB 471 SGQIKLETTQVSGNYGTON-SQQT-ITENGSWNSVISQIDSVSASIID-TGSQTFER 527
QY 504 RIAAVNPSPLETTKPDMTLKEALKIAPGFNPNGLQY-QGKDITE--PDFNFDQOTSQ 560
DB 528 RVAAKEQGNPEDKT-PEITIGEAIKKAFSATK-NGELLYFNGIPIDESCVELIFDDNTSE 585
QY 561 NIKNQLAEIATNIYVLDKIKNKAKNNILIRDKRF--HYDR-NNIAVGADESVEKAEHR 617
DB 586 IIEQQLKYLDDKKLYNV-----KLERGNLILKVPSTFTNDEFENPP--ASWNIDTKNQ 639
QY 618 EVINSSTEGT-----LLNIDKDIRKILSGY-----IVEIETEGEKEVINDB 658
DB 640 DGLQSVANKLSEGTIIIPMSKLPKYKYVFSGVSKDPSTNSITVNIKSQKTDYLV 699
QY 659 RYDMLNIS-----SLRODGTFFDFKYNKDLPLYISNPNKYVNVAVTKENTIINP 696
DB 700 EKDYTKFSYEFETGKDSDBITLTSNGVIFLDNLISITELNSTPEILKEPIKVP 759
QY 697 -----NVYAVTKENTIINPSENGTSTNGI 721
DB 760 ILDAHKKYADIKLDT-----NTGNTYIDGI 785

RESULT 3
I39933
cryptic protein - Bacillus anthracis
C:Species: Bacillus anthracis
C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999
C:Accession: I39933
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppa, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr
A:Reference number: I39933; MUID:89172073; PMID:3148491
A:Accession: I39933
A>Status: preliminary; translated from GB/EMBL/DDJ
A:Molecule type: DNA
A:Residues: 1-192 <RES>
A:Cross-references: GB:M22589; NID:g143280; PIDN:AAA22636.1; PID:g143281

Query Match 6.2%; Score 235.5; DB 2; Length 192;
Best Local Similarity 34.6%; Pred. No. 2.4e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;
```


QY 587 MNILIRDKRPHYDRNNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEI 646
Db 1 MNILVRDP-YHYDNGNIVGVDDSYLKNAYKQILNWSGVSINLDEVDNQALSGYMLQI 59
QY 647 EDETE-----GLKEVINDRYDMLNLSLRQDGKTFIDFKYNDKLPYISNPN 693
Db 60 KKPNSHLNTPVTITILAGKDSGVGELYRVL-----DGAGFLDNKFDENWRSIV-DPG 112
QY 694 YKVNVAVTKEN-TIINPSENGDTSNGIKKILIFSKGYEI 734
Db 113 DDVYVAVTKEDFNATVDRDENGNI-A-NKLKNTLVLSGKIKEI 153

RESULT 4
G59104
hypoetical protein pXO1-111 - Bacillus anthracis virulence plasmid pXO1
C:Species: Bacillus anthracis
C>Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-May-2000
C:Accession: G59104
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
J. Bacteriol. 181, 6509-6515, 1999
A:Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbored
A:Reference number: A59091; MUID:199445483; PMID:10515943
A:Accession: G59104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <OKI>
A:Cross-references: GB:AF065404; NID:94894216; PID:AA032415.1; PID:94894327
A:Experimental source: strain Sterne
A:Note: similar to hypoetical protein in the protective antigen domain; ypa, plasmid p
C:Genetics:
A:Gene: pXO1-111
A:Genome: plasmid

Query Match 6.2%; Score 235.5; DB 2; Length 204;
Best Local Similarity 34.6%; Pred. No. 2.6e-05;
Matches 56; Conservative 36; Mismatches 4; Indels 23; Gaps 6;

QY 587 MNILIRDKRPHYDRNNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEI 646
Db 1 MNILVRDP-YHYDNGNIVGVDDSYLKNAYKQILNWSGVSINLDEVDNQALSGYMLQI 59
QY 647 EDETE-----GLKEVINDRYDMLNLSLRQDGKTFIDFKYNDKLPYISNPN 693
Db 60 KKPNSHLNTPVTITILAGKDSGVGELYRVL-----DGTGLDNKFDENWRSIV-DPG 112
QY 694 YKVNVAVTKEN-TIINPSENGDTSNGIKKILIFSKGYEI 734
Db 113 DDVYVAVTKEDFNATVDRDENGNI-A-NKLKNTLVLSGKIKEI 153

RESULT 5
F82885
hypoetical protein UV482 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82885
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A:Reference number: A82870
A:Accession: F82885
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-468 <GLA>
A:Cross-references: GB:AE002145; GB:AF222894; NID:96899476; PID:AAF30894.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: UV482
A:Genetic code: SGC3

Query Match 5.4%; Score 202; DB 2; Length 468;
Best Local Similarity 21.3%; Pred. No. 0.13;

Matches 174; Conservative 131; Mismatches 325; Indels 186; Gaps 40;
QY 19 LLGYFSDLNQFQPMVVTSTTODLSIPSELENIPSENOYFQSAIWSGFIKVKSDYVT 78
Db 3699 LVDVYILD-NIHQNI DETRKIFKDHV-SKETEINPGVTMISKHGWNKSPDTDTANFEFK 3756
QY 79 FATSADNHTVMWVDDOEVIKASNSKIRLEKRLVQIKIQVORENPTKGLDFKLYWTD 138
Db 3757 IETQ-----DDNDVLNIDATVFKDEHNNIKQIVRIKEN-----ND 3795
QY 139 SONKEVISEDNLQLPKQKSSN-----SRKKRSTAGPTVDPDR----- 179
Db 3796 WLIKQI---DNLN-PETKYKLENIELSKPLKTHHTNLSVINDKENISLITETGNPVLKV 3851
QY 180-----NGIDPSLEVEGYTVGVKVRFLSPWISNIEHKKGLTKYKSPKKNSTASPYSD 235
Db 3852 IQONTDINTDQGTINTVLSGVNSK-YNGRQIKVYKDNNNVYESS---LITLQKGNK 3907
QY 236 FEKVTGRIDQVSPPEARHPVLAAPVHVDMENIILSKN-EDOSTQNTDSET-RTSKNT 293
Db 3908 YQLLSNLNSN-----REYFEKIEIHHSNTNNFEDLEKLGVSNTFITOTKNT 3957
QY 294 STSRTHTS-EVHGNAEVHASF-----FDGGSVSAGFS-----NSNS---STVAID 335
Db 3958 TVQWINDSSATVTRGVNFNFKIKSEDKILENNQVVAWFAPKETTIRDTNTMLQYTRPLX 4017
QY 336 HSLSLAGERTWAETMGLNT---ADTARLANATRYVNTGTAPIYVNLPTTSLVLGKQOTLA 392
Db 4018 DVTSDPKEGTWADLNSVNFKEETTYKLVQFVNKPTKAKNNINNSNVILDTNNSI 4077
QY 393-----TIKADENQLSOLAPNNYPSKMLAPALN-AOKDASSTPITWNY--NOFLELE 443
Db 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQTINFTLSGVKKSVMGKKIKLSYKSNDSSES 4137
QY 444 KTQLRLDQVYGNATVNFENGR-----VRVDGSKNWEVLPOIQETTARIENGKD 497
Db 4138 HTNEVLIESKTOYINLLNKLKRNRYTTLIDVKLIDNNVSDFPKEGNTNNSFITTRTSA 4197
QY 498 LNLVERRIAAVNPSPLETT-----KPDMLTKEA-----IKIAGFNEPNENGLQ 541
Db 4198 INVLTIEISNRASNTLKSITIKINLNDPQVLRDQQTIVVGNKNQAMGFITVSGNIK 4257
QY 542 YQKDKTEPFDNFQDQTSQNIK-NQLAELNATYIVLDKIKLNAKQNTILIRDKRPHYD- 599
Db 4258 YLTATLVDLNFN-DKVNIVNISFNPKPSAAEN-----IGDKSNII-----YNNDS 4304
QY 600-----RNNTAVGA---DSVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEI EDETE 650
Db 4305 IPKLEINNDIIVNGPINKKEIVVKNQKQ--NNIDVDLGLQINPKIAHNL-R-FLAKPKSTN 4361
QY 651 GLKEVIND--RYDMLNLSL-RODGKTFIDF-----KKY----- 681
Db 4362-----NDIETVINGSSLVWNGDKTSIRFTLNNLKANKLYSLVDVYVYLVNNSNTIVE 4415
QY 682 NDKLPLYISNPKYKAVYAVTKENTINPSENGDTS 717
Db 4416 SNKLP-KLNNINYOIK---INKSHTII--SKNGEWS 4445

RESULT 6
T28676
rhostry protein - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28676; A45521
R:Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem. Parasitol. 76, 329-332, 1996
A:Title: Comparison of two members of a multigene family coding for high-molecular mass
A:Reference number: Z20507; MUID:97077455; PMID:8920022
A:Accession: T28676
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2401 <SIN>

A;Cross-references: EMBL:U36927; NID:gl041784; PID:gl041785; PIDN:AB41263.1
R;Reen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A;Title: Identification of the gene for a Plasmodium yoelii throptry protein. Multiple co
A;Reference number: A45521; MUID:91101660; PMID:2270106
A;Accession: A45521
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 2260-2401 <KEP>
A;Cross-references: GB:IM34281

Query Match 5.1%; Score 190.5; DB 2; Length 2401;
Best Local Similarity 21.7%; Pred. No. 0.19;
Matches 182; Conservative 119; Mismatches 316; Indels 221; Gaps 43;
QY 1 EYKQENRLINSESSQGLGYFSDLN-----PPAPVVTSSITGDLST 45
DB 402 EYAKENVOLNVYKSNLEIKHYNDOINIDNIKEAKONYDQFKEHMTIPPNEKMYQK 461
QY 46 PSELENIPSENOYQSAT--WSGPIKV---KKSDEYTFATSAONHVMTWDDQEV---I 97
DB 462 PSIEIKMKDE--FLSKVKNYDFDKVYKVESEHKNFTELTNKIITEVSDIEIKYE 518
QY 98 NKASNSKIRLEKGLYQIKIYQRENTEKGLDFKLYWTSQNKKEVISSDNILQELK 157
DB 519 NKFNDSKSLNETKK--SIEEYQNLTKKYDD---YIKVCLNTNELITNCHNKQTLK 573
QY 158 QKSSNRK--KRSTAGPTVPDRDGDIPD-----SLEVEGYTVDVKNKRTFLSP 205
DB 574 DKLQNKIKTKETNSIDKTYTDKFENILTDKTELETKTGLSNHESNKKELLTPYD 633
QY 206 WTSNTHKEKGLPKYKSSPKWSTADPSYDFKVTGRIDKNVSPPEARHPLVAAPVHVD 265
DB 634 LKANLGNKENMLYQFNEK---BKAVEDIKKQVNDIKVSN-----IBIT 677
QY 266 MENILSKNEDQSTQNTDSETRTISKNTSTSTHTSEVHGAEVHASFDDIGSVSAGFS 325
DB 678 IYTSIYNINED---TENE---IGKSELNTKVLK-----KVKANVT 713
QY 326 NNSNSTVAI--DHSLSLAGERTWAETMGLNTADTARLANIRYVNTGTAPIYVLTSLV 384
DB 714 NLNEIKELKDYDFQFGK-----EKNIKYPDEN--KIKNDIDTLNQK 754
QY 385 LGKN-QTLATIKAD-ENQLSQI-----LAPN---NYYP---SKNLAPALNAQK-- 425
DB 755 IDKSIETLTETIKNSNENHIDEIKQIDKUKVFNKTMFNEDEPKIEKKIENIVEKIDKK 814
QY 426 -----DASSTPITMYNOFLEKTKQLRLTDQVYGNIAATYFNGRVRVDTGSNWE 479
DB 815 NIYKEIDKLLNE-SKIENTDKTSLEKLNLSYKSLGNLPLQDIDEEKKAEHTIKAME 874
QY 480 V-----LPQIETARIIFNGKDLNL-----VERIAAVNPSDP-----LETT---K 518
DB 875 AYIDDLNITKKKSQEI---EKEMNINMDIKMDIHKEMKALNTSHDOYKIYHTTSKHBEK 931
QY 519 PMTLKEALKIAPGFNEP---NGNLQYQGDITFDFNFDQOTSQNIKNQLAELNATNIY 575
DB 932 ISPIRKNLSKLIQDFSEESYINDIKKELEKNVLE-----SQNNNTDINQYLSKIE--NIY 984
QY 576 TVLDKTKLNAKXNMLIRDKRF--HYDRNNAIAGADESVVKEAHREVNINSTEGLLNIDK 633
DB 985 NIL---KLAKIKKIIDKVKYEYTDIEKKN-----KKINAELSNS-----1020
QY 634 DIRKILSGYIETEDTEGLKE-----VINDRY---DMLNSSLRQDQKTFIDPKKYND 683
DB 1021 --EKI-----ITQKENSLLKECOSKIKSTIDDNVYSECIKNTNL-----KTVIVNEKN-- 1069
QY 684 KLPYTSN-PNKKVNV-----YAVTKENTINPSENGDTSTN--GIKKILIFSKK 730
DB 1070 -INTYFNAEYVQNVSLNFNFT-EMADTKSQYLINIKKNGTNTDYNIKELKEHKKK 1126

RESULT 7

B64635
toxin-like outer membrane protein HP0922 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C;Accession: B64635
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, I
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: B64635
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2529 <TOW>
A;Cross-references: GB:AE000602; GB:AE000511; NID:g2314060; PIDN:AD07969.1; PID:g231410

Query Match 5.0%; Score 189.5; DB 2; Length 2529;
Best Local Similarity 21.0%; Pred. No. 0.23;
Matches 158; Conservative 92; Mismatches 264; Indels 239; Gaps 35;
QY 123 ENTEKGLDFKLYWTSQNK---KEVISSDNILQELKQKSSNRKSTKSTAGPTVPDR 178
DB 729 QNPASVSWGVTL--QNKTSNKGIIYID-----PNLSGQSGQSGNTLSTYANLF--- 779
QY 179 DNDGIPDSLEVEGYTVDVKNKRTFLSP---WI-----SNHEKKGITKYK 220
DB 780 ---GRSFSVNIQNGTLIIGNNTESVNSGLIWHGFGYITGTFSAAIY---LTNMF 832
QY 221 SSEKYST-----ASD-----PYSDFEKVTGRIDKNVSPPEARHPLVAAPIV 262
DB 833 KTGEVNSDGGGANTTFKASDNITWDGLNYNDAAETVTKIOTGAS---CHSYATFDALN 889
QY 263 HVDENIILSKNEDQSTQNTDSETRTISKNTSTSTHTSEVHGAEVHASFDDIGSVSA 322
DB 890 NISVTNSSF-----DMTWKFSFSAKNISFS-----NASF-----S 921
QY 323 GFSNSNSTVA---IDHSLSLAGERTWAETMGLNTADTARLANIRYVNTGTAPIYVNL- 378
DB 922 GFTNPGOGSSVISANATNSLSFNSR-----LNGAVYNLQANSIIFN-NTQAVFNVLV 973
QY 379 -----PTSLVLGKNQTLATIKADENQLSQIILAPNNYPSKNIAPIAL-NAQKASS 429
DB 974 SRGTSNPNATQLLGNNTFLSS-----QSLNFGDITLQNNANITLGNKSQAAPK 1025
QY 430 TPTMANNQFLEKTKQLRLTDQVYGNIAATYFNGRVRVDTGSNWEVLPIQIETTA 489
DB 1026 NSULTDNNSLSDQNSVLNANNTSAFNQASLIYNGS-----QATFN 1069
QY 490 RIIFNGKDLNL-VERIAAVNPSDPLETTKPDMTLKEALKIAPGFNEPNGNLQYQGDIT 548
DB 1070 SLFPNGTSLNASSKLNASNASFSNNTT---INLDDSVLSASNTSSLANINFQASQA 1126
QY 549 EF-----DFNFDQOTSQNIKNQLA-----EL 569
DB 1127 DFGGNTIITDASTFDSASLSLNFNLTANGALNFGYTPSLTKALMSVSGQFVLGNNGDI 1186
QY 570 NATNIYTVLDKIKLNAKXNMLIRDKRF-----HYDRNNAIAG 606
DB 1187 NLSDI-NIFNITKSVTYNIIILNAQKGITGSGANGYEKILFYGMKTKQNTATYSDNNIQTW 1245
QY 607 A-----DESVVKEAHR-----EVIN--SSTEGLLNIDKDI-----RKILSGYIVEI 646
DB 1246 SFINPLNSSQIIQESIKNGDLTIEVLNPNASNTIFNIAPELYNTYQASKQNPYTSYDY 1305
QY 647 EDTGLKEVINDRYDMLNLSL-----RODKTFIDPKKYNDKLPV---ISNPNYK 695
DB 1306 SDNOA-----GTYLTSNIKGLFTPKGQTPQAPGYPPNPQLSSLNINPKGFSSENK 1360
QY 696 VNVAVTKENTIIIN--PSENGDTSTNGIKKIL 725
DB 1361 TLLGILSQNSATLKEMIESNQLDNIIN-INEVL 1392

RESULT 8

E90598
membrane nuclease, lipoprotein [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: E90598
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: E90598
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1125 <KUR>
A:Cross-references: GB:AL45566; PID:G14090108; PIDN:CAC13866.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU_6930
A:Genetic code: SGC3

Query Match 4.9%; Score 186.5; DB 2; Length 1125;
Best Local Similarity 19.1%; Pred. No. 0.1;
Matches 165; Conservative 162; Mismatches 330; Indels 207; Gaps 40;

```
QY 3 KQENLLNSESQGLLGYPFDLNFQAPMVVTSSTGDLSPSELENIPSENQYFQS 62
DB 120 ENDNSVNNKNSSSK-----NDEN-----LTLTKVRLGHVNV-LNQGNIPTKMLAISK 168
QY 63 AIWS-----GFTKV-----KKSDEYTFATSDNHNVTMVVDQ-----VI 97
DB 169 VILHNKLDVLGLTEITNENGVKTIINELNKSDDKNVIVSLKRGTTGSSGQEHVGII 228
QY 98 NKASNNKIRLE-----KGLYQIKIQYQRENTEKGL-----130
DB 229 YK---ENKLTLESFDDKEKSKGRFYENKLW---DDPFKKGQKIDFVRPPFGVKFTKNI 282
QY 131 --DFKLW---TDQNKKE---VTSSDNQLPELKQKSSNSRKRSTPSAGTVPDRDNGI 183
DB 283 KNDFTVFMHSDAFGYKEERGEISAKGYSGQGHKEVAEARLTKVEYFDSI-----DGV 337
QY 184 PDSLEVEGYT-----VDVKNKRTFLSP-W--ISNIHEK---K 214
DB 338 NNELFFMGDTNLIKNGEAKAFELLOSQVKSLLIKDVKENATSLAQRWGEYANHYDKIFYK 397
QY 215 GLTKYKSP--EKWSTASDPYSDPEKVTGRIDK-NVSPARHP-----LVAAYPI 261
DB 398 GDLKVENSGFYDLWKVPDDNLLNKEEFKTKVEYRSKKAQKQKGYSVLHAISDHTI 457
QY 262 VHDVMEIILSKNEDQSTQNTDSETRTISKNTSTRTSTHTSEVHGNAEYHASFIDIGGSVS 321
DB 458 VYTDL--ILLQKDD---QNKNSNK--DENSDSKQNLDPKPTNNEQONTQDDSKKIS 509
QY 322 AGFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVLPIT 381
DB 510 DASQNSNTTNEKQKLDQDE---SKQNAIKSQQNDQKSDNSLSSSKNDTPQSKESPQI 566
QY 382 SLVLGKQNTATIKADENQLSQILAPNNYFSPKSLAPIALNAQKASSTPITMNYQPLE 441
DB 567 NPNLENNQEIHSNNGENDDSKEQNTSNSRQTKN---DLRSQKQNTL-----611
QY 442 LEKTKQRLDLDQVYGNIAVNFNGVR---VDTGNNWSEVLPQQTETARIFNGKOL 498
DB 612 ---TKNPSNSNNTVETKNETQNNNSSTKKDEIDTSA-----KTQDSTNSNLKNEKT 661
QY 499 NLVERRI-----AANVPSDPLETKPDMTLKEALKI--AFGNENPNGLQYQCKDITE 549
DB 662 NQVETKNTESNNSNINKQENSTKKEEISKSESNNVNSNNTKQENIDNKKKEISK 721
QY 550 FDFNPDQOTSQNIKNQ-LAEALNATNIYTVLDKIKLNKQNM-----ILIRKRRHYDRANI 603
DB 722 SESNVNNSNTNQETPETNESQNNVIGKNPNNSLNQNAIDVSAKKVKGIGMYNINE 781
```

RESULT 9

S05603
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (S
N:Alternate names: gp195 surface antigen
C:Species: Plasmodium falciparum
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jun-2000
C:Accession: S05603; S04850

R:Myler, P.J.
submitted to the EMBL Data Library, April 1989

A:Reference number: S05603
A:Accession: S05603
A:Molecule type: mRNA
A:Residues: 1-1639 <MYL>
A:Cross-references: EMBL:X15063; NID:G9896; PIDN:CAA33163.1; PID:G9897

R:Myler, P.J.
Nucleic Acids Res. 17, 5401, 1989

A:Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plas
A:Reference number: S04850; MUID:89345116; PMID:2668887
A:Accession: S04850
A:Molecule type: mRNA
A:Residues: 1504-1639 <MYL2>
A:Cross-references: EMBL:X15063

C:Superfamily: major merozoite surface antigen

C:Keywords: Glycoprotein; merozoite; surface antigen

F:1-19/Domain: signal sequence #status predicted <SIG>

F:120-1639/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 4.9%; Score 184.5; DB 2; Length 1639;
Best Local Similarity 20.5%; Pred. No. 0.22;
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

```
QY 5 ENRL-LNESESSQGLLGYPFDLNFQAPMVVTSSTGDLSPSELENIPSENQYFQSA 63
DB 927 ENILSLGKNKIYQELIGQKSE-NF-----YEKILKDSDTFYNE 965
QY 64 IWSGFIKVKSDYEYTFATSDNHNVTMVVDQEVINKASNS-----NKIRLEKGRLY- 114
DB 966 SFTNFVKSKADD-----INSLNDESKKKLEEDINKLKTQLQSLFDLYNKYKLERLFD 1020
QY 115 -----QIKIQYQRENTEKGLDPKLYWTDSONKKEVSSDNQLQPELKQKSSNRKKRS 168
DB 1021 KKKTVGKYKMQIKKLTLLKEQLESL--NSLNNPKHVL--QNFVSFFNKKKEAEIAETEN 1076
QY 169 TSAGTVVPRDNDGTPDLSLEVEGYTVQVKNKRTFLSPWISNIHEKKGITKY---KSSP-- 223
DB 1077 T-----LENTKILKHY-----KGLVKYNGESSPLK 1103
QY 224 ---EKWSTASDPYSDPE-----KVTGRIDKNVSPAR-----HPLVAAYPIVHVD 265
DB 1104 TLSESIQTEDNYASLENFKVLKLEGLKQNLKONLEKKLSYLSGLHLLIA-----E 1156
QY 266 MENIILSKNEDQSTQNTDSETRTISKNTSTRTSTHTSEVHGNAEYHASFIDIGGSVAGFS 325
DB 1157 LKEVTKNKN---YTGNSPSENNT-----DVNNALSYKKLPQGTDTVATVS 1200
QY 326 NNSSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVLPIT---- 381
DB 1201 ESGSDTLQSOPKPKASTHVGRES---NTITTSQ-NVDDEVDVVIIVIFGESEBDYDDL 1256
```

382 SLVGGKQTLATIKADENQLSOLAPNNYPSKNLAPIA---LNAQKDSASTPITMNYN 437
1257 GQVVTGEAVTPSVI---DNILSKI---ENEVEVLKPLAGYRSLKKQLENNVMTNVN 1310
438 ---OFLSEKTKQLRLDQV-YGNIAI-----YFNENGRVRVDTGNSWSEVL 481
1311 VKDIILNRPENKFNKV-LESGLIPYKDLTSSNVVVKDPYKFLNKEKRDKFLSSYNVYK 1369
482 POIQETIARIIFNGKDLNIVERRIAAVNSDPL-----ETTKPDM-TLKEALKIARGF 533
1370 DSID-----TDINFA-----NDVLGYKYLSEKYSKDLDSIKKYINDKQGE 1410
534 NEPN---GNLOVQKQDITE-----DFNFDQOTSQNIKNQLAELNATNIYT 576
1411 NEKILPFLNIEETLYKTVNDKIDLFVHLEAKVLNHYEK---SNVEVKIKELN--YLKT 1465
577 VLDKIKLNAAMN--ILIRDKRPHYDRNN-----IAGV-ADBSVVEKAHREVINSTEGLL 628
1466 IQDKLADFKNMNFVGIADLTDYNNHLLTKFLSTGMVFENLAKTVLSNLLDGNLQW-M 1524
629 LNTDKD--IRKILSGYIVIEDETEGLKEVINDRYDMLNITSSLRQDGKTFIDPKYNDKLP 686
1525 LNIHQCVKVK-----QCPQNSGCFRHLDE-----RECKCLNLYKQEGDKC- 1566
687 LYISNPNYKVNVAVTKENTIIINPSENG-----DTSTNGIKKI 724
1567 --VENENPTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1599

RESULT 10
B89921
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Ji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89921
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <KUR>
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BA842527.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebha

Query Match 4.9%; Score 183; DB 2; Length 6713;
Best Local Similarity 20.4%; Pred. No. 2.1; Indels 256; Gaps 43;
Matches 172; Conservative 125; Mismatches 125; Indels 256; Gaps 43;

9 LNESSSSQGLLYFSDNFQAPMVVTS-----STTGDLISIPSELENIPSENQYFQSA 63
2468 VRQAKSDAKNLG-TLTHLNAQKQDLTSQIEGATTGVNVSVKTKAQDLQDGMQRLESA 2526

64 TWSGFIKKYKSDXY-----TPTASDHNHTVWVDDQVINKASNSKIRLEKGLYQIKI 118
2527 I-ANKDQTKASENYIDADPTKTKTAPDNTAI---QAESYLNKDHGNTKDK-----QAVEQ 2576

119 QYQRENPTKGL--DFKLYWTDSONKKEVISSDNLQLPE---LKQKSNRKKKSTSGAP 173
2577 AIQSVTSTENALNGDANLQCAKTEATQAIIDLNTQLNTQKTKALQOVNNAQR----- 2628

174 TVPDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEK-----KGLTKYKSSPEK 225
2629 -----VSGVT-DLKNSATSLNAMDQLKQAIGDHDTIVAGNNTYVAGNPNASPDK 2672

226 WSTASDPYSDPEKVTGRIDKNVSPBARHPLVAAPYIVHVDMENITLSKNEDQSTQNTDSE 285
2673 QGAYTDYNAAKNIYNG--SPNVTNADVTATQV-----NNAETSLNGDTN 2719

286 TRTI---SKNTSTSRKTHSEVHGNAEVHASFFDIDGYSVA-----GFNSNSNSTVAIDHS 337
2720 LATAKQAKDALRQMTLSDAKQS-----ITGIDSAQTQVTGVQSVKDNATNLDNA 2771
338 L-----SLAGERTWAEATMGLNTADTARLANRYVNTGTAPIYVNLPTTS----- 382
2772 MNQLRNSIANKDEVKASQPYVDATDKQNA-----YNTAVTSAENIINATISQFTLPSAVT 2827
383 -----LVLGKQNTLATIK-----ADENOLSQILAENNYPSKNLAPIALNAQKQDAS 428
2828 QAAQVNTYKLTALNGAQLNANKQETAINRLSHL-----DIDQV-----YGNL 459
429 STPITMNYN-QFLELEKTKQLRL-----DIDQV-----YGNL 459
2872 NTQVTNAPNISTVNVQVTKAEQLDQAMERLINGIQDKQVQSVNFTDADPEKQTAYNNA 2931
460 AT-----YFNENGRVRVDTGNSWSEVLPOIQE--TTARIIFNGKDLNLRERRI--AAVNPS 511
2932 VTAENIINOANG-----TNANQSVAEALSTVTTTKQALNG-----DRKVTDAKNAN 2980
512 DPLET-----TKPDMTLKEALKIARGFNPNENGLQ--YQKQDITE 549
2981 QTLSTLDNLNNAQKGAVTGNINOAHVAVET--CAIQTAQELNATMGNLKNLNDKDTTL 3038
550 FDNFED-----QOTSQNIKNQLAELNATNIYTVLDKIKLNAKMMILIRDKRPHYD 599
3039 GSQNFADADPEKKVAYNEAVRVAENILKSTGTNV---PKQVVEAAMN-----QVN 3086
600 RNNTAVGADESVKAEHREVINSTEGL--LLNIDDKIRKILSGYIVETEDTGLKEVIN 657
3087 TTKAALNGTQNL--EKAKQHANTAIIDGLSHLTNAQKALQVQSTTVAEAQNSQKAN 3144
658 DRYDMLNLS-----LRQ--DGKTFIDFKYNDKLPYISNPNYKVNIV--AVTKENTII 708
3145 -----NVDAAMDKLRISADNATKQNYTD-----ASPN-KKDAYNNAVTTAQGII 3191
709 NPSEN 713
3192 DQTTN 3196

RESULT 11
JC6009
surface-located membrane protein lmp3 precursor - Mycoplasma hominis
C:Species: Mycoplasma hominis
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: JC6009
R:Badefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.
J. Bacteriol. 178, 2775-2784, 1996
A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene syate
A:Reference number: JC6009; MUID:96213016; PMID:8631664
A:Accession: JC6009
A:Molecule type: DNA
A:Residues: 1-1302 <LAD>
A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336
C:Genetics:
A:Gene: lmp3
A:Genetic code: SGCG
C:Superfamily: surface-located membrane protein lmp3; tetratricopeptide repeat homology
F:Keywords: duplication; membrane protein
F:24/Domain: signal sequence status predicted <SIG>
F:25-1302/Product: surface-located membrane protein lmp3 #status predicted <MAT>
F:957-992/Domain: tetratricopeptide repeat homology <TTL>
F:993-1026/Domain: tetratricopeptide repeat homology <TT2>
F:1089-1120/Domain: tetratricopeptide repeat homology <TT3>
F:1154-1190/Domain: tetratricopeptide repeat homology <TT4>

Query Match 4.8%; Score 181; DB 1; Length 1302;
Best Local Similarity 17.9%; Pred. No. 0.23;
Matches 137; Conservative 141; Mismatches 285; Indels 204; Gaps 28;

1 EVKQENRLNSESQGLLYFSDNFQAPMVVTSSTTGDLISIPSELENIPSENQYF 60

Db 476 QVDEANKSIKQALNLIKANTLLPQLNDNDSEIVKAKESLNABEITWANKAVNQDNASM 535
Qy 61 QSAIWSGFIKVKSDSEYTFATSAADNHVTWVDDQEVINKASNNKIRLEKGR--LYQIKI 118
Db 536 QSA-----KSSLDKVKTKIQNLTEFNKDKDAKFKLEQTRKIDNFLT 579
Qy 119 QYQRENPTKGI-----DFKLWYTDSONKKEVISSDNLQLPELKQKSSNSR--KKRS 168
Db 580 DDVKNPNYATLVKDLTWAKDDKSVTKSSNKSEIIAAND-----ELQALDKAKVAKQOI 635
Qy 169 TSAGTVPDRNDGIPDSLEVEGYTV---DVKNKRTFLSPWISNIHEKGL----- 216
Db 636 DEANKSIKQELSDSITNANQLLNKLVDGDKDIQAKTELSQEIQSASQELNLNPTSMQS 695
Qy 217 -----TKYKSSPEKWSIASD--PYSDFEKTGTRIDKNVSPPEARPLVAAPVIVHDM 266
Db 696 AKES--DAKVTEITKKLETFNKKDVKFKLEKTRKIDDEFINTKNTNP----- 743
Qy 267 ENILSKNEDOSTQNTDSETRTISKNTSTSRHTSEVHGNAEVAHSPFDIGGSVSAGFSN 326
Db 744 -----NYSTLSELTSKEDSKSVTNSNKSIDI----- 771
Qy 327 SNSSTVAIDHSLSLAG--ERTWAEWGLNLTADTARLANIRVYNTGTAPIYVWLTSTLVL 385
Db 772 -ETANTEKQALAKANTKAQADNLAKSTKE--QLNNSISSANTLLAKLTD----- 819
Qy 386 GKNOQLATTIKADENOLSQLAPNNYPSKNLAPI--ALNAQKQDASSTPI-----TWNVNQF 439
Db 820 -KDNVIOQAKTELEK--EVQKANQAVASNNYASMQSAKSLDAKYTEITKLETFNKKDKD 876
Qy 440 LELEKTKQRLTDQVYGNIAATYFENGVRVDTGNSNWEVLVPOIQTETARIIIFNGKDLN 499
Db 877 VKFKELEQTRKIDDEF-----INTNKTNPDYSTLISELTSK----- 912
Qy 500 LVERRTAAVNPSPDLETTKPDMLKEALKIAPGFNEPNCNLYQCGKDIETDFNPDQOTS 559
Db 913 -RDKNSIITNSNKSIDETANTEKQALAKA-----NTDKQAD--NLARSTK 957
Qy 560 QNKKQALNATNIYTVLDKIKLNKAKMILIRKRFHYDRNNIAVGAD---ESVVKBAH 616
Db 958 EQLNKSISSAN-----TLAKLT-----DKNVTIQAQKTELEKEVQKAN 996
Qy 617 REVINSSTGLLNIDKDIRKILSGYVIEIDTEGLKEVINDR--YDMLNIISSLRDQKTF 675
Db 997 QAVASNTASM-----QSAKSSLDKAVTEI--TKLETFTNKKDVKFRELEQTRKIDDEF 1049
Qy 676 IDPKYNDKPLVYISNPKNVYAVT-----KENTIIINPSENGDTST 718
Db 1050 INTNK-----TWPNYSTLISELTSKRDSKNSITNSNKSIDET 1087
RESULT 12
T28677
rhoptry protein - Plasmodium yoelii
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28677; C45521
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 65, 171-177, 1994
A:Title: A gene coding for a high molecular mass rhoptry protein of Plasmodium yoelii.
A:Reference number: Z20508; MUID:95021522; PMID:7935623
A:Accession: T28677
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-2269 <KEE>
A:CROSS-references: EMBL:L27838; NID:g457145; PID:g457146; PIDN:AAA21304.1
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A:Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co
A:Reference number: A45521; MUID:91101660; PMID:2270106
A:Accession: C45521
A>Status: preliminary

A:Molecule type: DNA
A:Residues: 2131-2269 <KE2>
A:CROSS-references: GB:M34283
Query Match 4.8%; Score 181; DB 2; Length 2269;
Best Local Similarity 20.5%; Pred. No. 0.53;
Matches 161; Conservative 121; Mismatches 301; Indels 204; Gaps 35;
Qy 1 EVKQENRLNSESQSGLLGYVFDLNPQAPWVVT---SSTGDLSPSS-----L 50
Db 587 KVENVTNLNK-K--EKLKHYDFDFGKEGNIKYTDKIKKINDDIMAVSQIQDHINGL 643
Qy 51 ENTPSNQVQSAIWSGFIKVKSDSEYTFATSAADNHVTWVDDQEVINKASNNKIRLEK 110
Db 644 DDIQKSESIVSEMKEQINKLEKVSNTETI--SNDNVEGIGKKQOIIVTIDKKNKIYEI 701
Qy 111 GRLYQIKIYQRENPT--EKGDLFKLWYTDSONKKEVISSDNLQLPELKQKSSNSRKRST 169
Db 702 NKLSISIKEDKNTSLEKVKDINLSY--GQNLGHLFLE---QIDEEKKAENTIK--- 752
Qy 170 SAGTVPDRDN-----DGIPDSLEVEGYTVVKNKRTFLSPWISNTHKGLTKYKSSPE 224
Db 753 SMEAYIDDLNKKKQSEIETEMDIK---MDINKEMEALK--ISHDDDKKCHDKSNHKE 807
Qy 225 KMTASDP-----YSDFEKVTGRIDKNVSPPEARPLVAAPVIVHDMENIILSKN 274
Db 808 NISDIYDKSKIIQDPSRESIDINDIKKLOKNVSESONH-----NSDINQCL---N 855
Qy 275 EDOSTQNTDSETRTISKNTSTSRHTSEVHGNAEVAHSPFDIGGSVSAGFSNSSTVAI 334
Db 856 EVANIYNI-LKLNKIKKIIDKVKYETSEIEKNNK-----NINDELNNSEKVIKKI 904
Qy 335 DHSLSLAGEWTWAEWGLNLTADTARLANIRVYNTGTAPIYVWLTSTLVLGKNQTLATI 394
Db 905 EGLSLKCEKSKTN---STLDDKDIDECIKNTN-----VLKKN----- 939
Qy 395 KADENQLSQLAFNNYPSKNLAPIALNAQKQDASSTFITMNVNQFLELEKTKQLRLDQ 454
Db 940 -----ILNEE-----TNIT---NHFKNAEE----- 956
Qy 455 YGNIATYFENGVRVDTGNSNWEVLVPOIQTETARIIIFNGKDLNVERRIAANVPSDPL 514
Db 957 -YKNIVLSNFN---IEMADNKSQVILEIKONGT---NDHDYNIKELK-SHKDSNGY 1007
Qy 515 ETTKPDMLKEALKIAPGFNEPNCNLYQCGK-----ITEFDNPDQOTSONIKN 564
Db 1008 K-TEADQNKAKIQK-----NKELEQYKEEVTVLNNKYAVELKKNKFD-KTKDMSKQ 1057
Qy 565 QLAELNATNIYTVLDKIKLNKAKMILIRKRFHYDRNNIAVGADSESVVKBAHRE 618
Db 1058 IIEIKDAHNYCTLESKSEKQKNEIKNEKIHIEDEVANNNDKSKAKITSIKVSEPPPTK 1117
Qy 619 VIN---SSTEGLLNIDKDIRKILSGYVIEIDT-----EGLKEVINDRYDMLNI 665
Db 1118 IIKINEIRTKSDCCLEKTNDEKQISNLSIDTQETKLTENGKQLKLEEL-----L 1168
Qy 666 SSLRQDQKTFIDFKKYNDKPLVYISN-----PNYKNVYAVTKENTIIINPSENGDT 716
Db 1169 ESLKKQKKNIEDQKDELDEVNSKIKNIENVTNCHKNYEIGI--VEKINEIAKTNKQIE 1226
Qy 717 STNGIKK 723
Db 1227 STKELIK 1233
RESULT 13
A86827
Hypothetical protein yvfg [imported] - Lactococcus lactis subsp. lactis (strain ILL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: A86827
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrh
Genome Res. 11, 731-753, 2001

RESULT 14

```

Qy 540 -----LOYGKDI-TBFDEFD-----QOTSONIKNOELNATNIYTV----- 577
Db 850 VIGDASGKNTSEY-SKDIGVYGINSDALYNGSNQISPSNDFENGLTNSFIYFWL 908
Qy 578 --LDKIKLNAXM-----NILRD-----KRPFY----- 598
Db 909 RNLGKDTIKSLIGSKEDNCGWEIYFQDTGLVFNMDNSNGEKNIYLSDVNSNWHYITI 968
Qy 599 --DR--NNIAGDAESV-KEAHREVINSSTEGLLNIDKDIRKILSYIYEIEDTGLK 653
Db 969 SVDRLUKEQLLFIIDNLVANSIKELIYNSNIISLSENNPSYIEGLTILNKPTTS-Q 1027
Qy 654 EVINDRYMLNLSLRQDKTFIDFKYNDKLPYISNPYKNV---YVAVTKENTI--- 707
Db 1028 EYLSNVEFLANSYIRDSNEERLE---YNKTYQLY---NVVFSKPICEVKQNNIYLT 1080
Qy 708 INPSENGDSTNGIKKILIFSKGY 732
Db 1081 INNTNMLQASKFKLLSINPNKQY 1105

RESULT 15
C97033
uncharacterized protein, probably surface-located [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: C97033
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97033
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1227 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79054.1; PID:g15023995; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1080

Query Match 4.7%; Score 179; DB 2; Length 1227;
Best Local Similarity 18.9%; Pred. No. 0.27;
Matches 166; Conservative 136; Mismatches 288; Indels 288; Gaps 44;

Qy 37 SBTGDLSPSELENIPSENQYFQSAIW-SGFIKVKSDFTFATSADNHVTKWVDQE 95
Db 14 SVTTSVLLSKPAIKAYAADNSVYKSSISNSNEINVKGK-----VQSE 58

Qy 96 V-INKASNGKIRLEKGRLYQIKIQVRENPTK-----GLDFKLYWTDSQNKKEVI 146
Db 59 VNVKNKNDNS-KVSSSNENQVS---NKNSNPKVSSSSSEIQSINKNVNLQVQNNKSVL 114

Qy 147 SSDNLQLPKOKSNSRKK-----RSTSGTVPDRNDGIPDSLEVEGYTVDVKNRT 201
Db 115 AASNVDDEVKINGNVQTSYIAIGETKVKPDTLLIINKAIVDARSAG--TDLSEVEI 172

Qy 202 F-----LSPWISNI-----HEKKG-LTKVK 220
Db 173 YDIVSQTAEIAQAFRNINDGVANVSDYTLGATFVNDANLDSVKNKYFHKRYATVTKF 232

Qy 221 SGPWKSTA-----SDPYSDFEKTGTGRIDKNVSPPEARHPLVAA 258
Db 233 DVATKTSNALKNNNGQGETDYTALEVSGVTPYLDL-----VNKNIVKEKQ----- 280

Qy 259 YPIVHVDMENILSKNEDQS-TQNTDSETRTISK-NTSRTHTSEVHGNAEVHASFEDI 316
Db 281 -----NKGRLTIITEISDSAASTIARINTALDN-----MDA 311

Qy 317 GGSVSAGFSNSSTVAIDHSL---SLAGERTWAETM-----GLNTADTARLNANI--- 364
Db 312 GVATLEDYQAIAGANNVPQLHVADVNSLANDQRWGDVSEALDGINTIMTYINNINSVGTE 371

```

```

Qy 365 -RYVNTGTAPI-----YVNLPTTSL--VLGKQNTLATIKADENQLSQILAPNNY----- 411
Db 372 DDYINSHAVDSNEGNIIDYDILNANIIEKKTAKGQDL-TIPEVANVYKEVKTLDDFFNHAA 430
Qy 412 -----PSKULAPIALNAQKASSTPTMTNTNQFLELEKTKQLRLDQV----- 455
Db 431 AGOTTLQDYKNVDENPAQVQDDVAT-----LSDMLKTRDCKTLKALQDKIDSIILSLKQ 484
Qy 456 ---YGNATYTFENGVRVDTGSNNWSEVLPOIQETTARI-IFNGKDLNLVERRIAAVNP 510
Db 485 INSGIGNIDDYS-----KLQTEAVDASKLEAVNDDIKIKADKGEDLTIQEIRDSVKKT 538
Qy 511 SDPLETT---KPDWTLKBAIKIAP-GFNEPKNLQYQKQITEFDNFDDQOTSNIK-- 563
Db 539 IDYINSTSVSKGDSVSYITIGIDGVTEI--NIEFVNERIKESGITI---TIENIKV 593
Qy 564 ----NOLAELN---ATNIYTVLD-----KIKLNKMLILRDKEPHVD 599
Db 594 IEFIVQLSEVYRIVTGVGTVDVYKTLGINNVNDNNIYINAEKKNKDKVKIQDITRVD 653
Qy 600 R--NNI-----AVGADES-----VKEAHREVINSSTEGLLNIDKDIRKILSYIVE 645
Db 654 NTINNIDVINKIGAGDAVLSDYFNIGITDYOQDILDYVNAVADLKIQNYKDVDDIIEVREAK 713
Qy 646 IEDTEGLK-----EVINDRYDMLNLSLRQDKGTFIDFKYNDKLPY-----ISNENYK- 695
Db 714 ISSYEALMRINIGEAVTDDFKALGLTDI-----NDGLLELYATTDLQNKNYKT 760
Qy 696 -----VNVY-AVTKENTIIINPSENGDSTNGI 721
Db 761 ADEVIARVQAQIEIYRALMQIN--LGKATTADVTNLGI 796

Search completed: May 3, 2004, 19:41:23
Job time : 17.0855 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 8.2539 Seconds
(without alignments)
4636.784 Million cell updates/sec

Title: US-09-848-909A-13

Perfect score: 3772
Sequence: 1 EVKQENLLNESSSOGLL.....TSTNGIKKILIFSKKGYEIG 735

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3750	99.4	764	1 PAG_BACAN	P13423 bacillus an
2	235.5	6.2	204	1 YBI_BACAN	P34422 bacillus an
3	184.5	4.9	1630	1 MBP1_PLAEPK	P04932 plasmodium
4	184.5	4.9	1639	1 MBP1_PLAEPK	P04933 plasmodium
5	169	4.5	1658	1 YM67_YEAST	Q03661 saccharomyc
6	168	4.5	1000	1 S155_YEAST	P43612 saccharomyc
7	166	4.4	1276	1 BXD_CLOBO	P19321 clostridium
8	165	4.4	1803	1 YUL3_YEAST	P47024 saccharomyc
9	164	4.3	1957	1 SFOF_SCHPO	Q10411 schizosacch
10	162	4.3	1385	1 FAT1_SCHPO	O13735 schizosacch
11	161.5	4.3	1230	1 SMC3_YEAST	P47037 saccharomyc
12	160.5	4.3	979	1 P115_MYCHR	P41508 mycoplasma
13	160.5	4.3	1487	1 MDS3_YEAST	P53094 saccharomyc
14	160.5	4.3	2334	1 WAPA_BACSU	Q07833 bacillus su
15	158.5	4.2	1167	1 CAGA_HELPJP	Q921t1 helicobacte
16	158.5	4.2	1233	1 YF16_YEAST	P43597 saccharomyc
17	158.5	4.2	2116	1 MYS2_DICDI	P08799 dictyosteli
18	158.5	4.2	2869	1 RBP1_PLAEB	Q00798 plasmodium
19	157.5	4.2	1024	1 RIP3_MOUSE	P97434 mus musculu
20	157.5	4.2	2867	1 RBP2_PLAEB	Q00799 plasmodium
21	156	4.1	918	1 YVJB_CABEL	P34487 caenorhabdi
22	155.5	4.1	1208	1 PCP1_SCHPO	Q92351 schizosacch
23	155	4.1	1037	1 KC04_YEAST	P25389 saccharomyc
24	155	4.1	1790	1 US01_YEAST	P25386 saccharomyc
25	155	4.1	1928	1 MYS1_YEAST	P08964 saccharomyc
26	154.5	4.1	1460	1 M159_YEAST	P40477 saccharomyc
27	153	4.1	1288	1 VACA_HELPJP	Q92kw5 helicobacte
28	152.5	4.0	1046	1 SBCC_LACLA	Q9cfz0 lactococcus
29	152	4.0	1029	1 RIP3_YEAT	Q98re6 rattus norv
30	151.5	4.0	1271	1 Y338_MYCGE	P47580 mycoplasma
31	151	4.0	1250	1 BXE_CLOBO	Q00496 clostridium
32	150.5	4.0	1225	1 Y309_MYCGE	P47551 mycoplasma
33	150.5	4.0	1358	1 SIR4_YEAST	P11978 saccharomyc

ALIGNMENTS

RESULT 1

```
PAG_BACAN
ID PAG_BACAN STANDARD; PRT; 764 AA.
AC P13423; Q9F5R7; Q9KH69; Q9ROU2;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins
DE translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].
GN PAG_A OR PAG OR PX01-110.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89172073; PubMed=3148491;
RA Welkos S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppla S.H.,
RA Schmidt J.J.;
RT "Sequence and analysis of the DNA encoding protective antigen of
RT Bacillus anthracis.";
RL Gene 69:287-300(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=28, 33, BA1024, and BA1035;
RX MEDLINE=99214082; PubMed=10197996;
RA Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;
RT "Genetic diversity in the protective antigen gene of Bacillus
RT anthracis.";
RL J. Bacteriol. 181:2358-2362(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=V770-NP1-R / ATCC 14185;
RX MEDLINE=20359347; PubMed=10899854;
RA Cohen S., Mendelson I., Altboum Z., Kobiler D., Elhanany E., Bino T.,
RA Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M.,
RA Kronman C., Velan B., Shafferman A.;
RT "Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus
RT anthracis spore vaccines protect against anthrax.";
RL Infect. Immun. 68:4549-4558(2000).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=99445483; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
RA Keim P., Kohler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
RA Martinez Y., Ricke D., Svensson R., Jackson P.J.;
RT "Sequence and organization of pX01, the large Bacillus anthracis
RT plasmid harboring the Anthrax toxin genes.";
RL J. Bacteriol. 181:6509-6515(1999).
RN [5]
RP DOMAINS.
RX MEDLINE=91332080; PubMed=1651334;
RA Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;
RT "The carboxyl-terminal end of protective antigen is required for
RT receptor binding and anthrax toxin activity.";
```

Q96173 homo sapien
P53968 saccharomyc
P38931 saccharomyc
P16466 proteus mir
Q00816 saccharomyc
P54290 rattus norv
P10844 clostridium
P29539 saccharomyc
P18494 saccharomyc
O14157 schizosacch
Q03213 saccharomyc
Q53654 staphylococ

RL J. Biol. Chem. 266:15493-15497(1991).
RN [6]
RP CHARACTERIZATION.
RX STRAIN=Stearne;
RA MEDLINE=94327640; PubMed=80511159;
RT Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;
RN "Anthrax protective antigen forms oligomers during intoxication of
RL mammalian cells".
RL J. Biol. Chem. 269:20607-20612(1994).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=21129592; PubMed=11207581;
RA Beauregard K.E., Collier R.J., Swanson J.A.;
RT "Proteolytic activation of receptor-bound anthrax protective antigen
RN on macrophages promotes its internalization".
RL Cell. Microbiol. 2:251-258(2000).
RN [8]
RP TOXIN REGULATION.
RX STRAIN=Weybridge;
RA MEDLINE=94131936; PubMed=8300513;
RT Koehler T.M., Dai Z., Kaufman-Yarbray M.;
RN "Regulation of the Bacillus anthracis protective antigen gene: CO2 and
RL a trans-acting element activate transcription from one of two
RN promoters".
RL J. Bacteriol. 176:586-595(1994).
RN [9]
RP MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.
RX STRAIN=Stearne;
RA MEDLINE=95050722; PubMed=7961869;
RT Singh Y., Klimpel K.R., Arora N., Sharma M., Leppla S.H.;
RN "The chymotrypsin-sensitive site, FFD315, in anthrax toxin protective
RL antigen is required for translocation of lethal factor".
RL J. Biol. Chem. 269:29039-29046(1994).
RN [10]
RP MUTAGENESIS OF DOMAIN 4 LOOPS.
RX STRAIN=Stearne;
RA MEDLINE=99185012; PubMed=10085028;
RT Varughese M., Teixeira A.V., Liu S., Leppla S.H.;
RN "Identification of a receptor-binding region within domain 4 of the
RL protective antigen component of anthrax toxin".
RL Infect. Immun. 67:1860-1865(1999).
RN [11]
RP MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.
RX STRAIN=Stearne;
RA MEDLINE=21092804; PubMed=11178978;
RT Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;
RN "Trp 346 and Leu 352 residues in protective antigen are required for
RL the expression of anthrax lethal toxin activity".
RL Biochem. Biophys. Res. Commun. 281:186-192(2001).
RN [12]
RP MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.
RX STRAIN=Stearne;
RA MEDLINE=21438996; PubMed=11554763;
RT Anuja N., Kumar P., Bhatnagar R.;
RN "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are
RL required for oligomerization of anthrax protective antigen".
RL Biochem. Biophys. Res. Commun. 287:542-549(2001).
RN [13]
RP MUTAGENESIS OF PRO-289.
RX STRAIN=Stearne;
RA MEDLINE=21255689; PubMed=11356563;
RT Khanna H., Chopra A.P., Arora N., Chaudhry A., Singh Y.;
RN "Role of residues constituting the 2beta1 strand of domain II in the
RL biological activity of anthrax protective antigen".
RL FEMS Microbiol. Lett. 199:27-31(2001).
RN [14]
RP MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.
RX MEDLINE=21125576; PubMed=11222612;
RA Mogridge J., Mouruz M., Collier R.J.;
RN "Involvement of domain 3 in oligomerization by the protective antigen
RL moiety of anthrax toxin".
RL J. Bacteriol. 183:2111-2116(2001).
RN [15]
RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.
RX MEDLINE=21269403; PubMed=11113126;
RA Sellman B.R., Nassi S., Collier R.J.;
RT "Point mutations in anthrax protective antigen that block
RN translocation".
RL J. Biol. Chem. 276:8371-8376(2001).
RN [16]
RP MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;
RX STRAIN=Stearne;
RA MEDLINE=22112896; PubMed=12117959;
RT Chauhan V., Bhatnagar R.;
RN "Identification of amino acid residues of anthrax protective antigen
RL involved in binding with lethal factor".
RL Infect. Immun. 70:4477-4484(2002).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=97192099; PubMed=9039918;
RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;
RT "Crystal structure of the anthrax toxin protective antigen".
RN Nature 385:833-838(1997).
RN [18]
RP REVIEW.
RX MEDLINE=21428689; PubMed=11544370;
RA Mock M., Fouet A.;
RT "Anthrax".
RL Annu. Rev. Microbiol. 55:647-671(2001).
CC -!- FUNCTION: One of the three proteins composing the anthrax toxin,
CC the agent which infects many mammalian species and that may cause
CC death. PA binds to a receptor (APR) in sensitive eukaryotic
CC cells, thereby facilitating the translocation of the enzymatic
CC toxin components, edema factor and lethal factor, across the
CC target cell membrane. PA associated with LF causes death when
CC injected, PA associated with EF produces edema. PA induces
CC immunity to infection with anthrax.
CC -!- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a
CC protective antigen (PA), a lethal factor (LF) and an edema factor
CC (EF). None of these is toxic by itself. PA+LF forms the lethal
CC toxin (lethal); PA+EF forms the edema toxin (edix). PA-63 forms
CC heptamers and this oligomerization is required for LF or EF
CC binding. Once activated, at low pH, the heptamer undergoes
CC conformational changes and converts from prepore to pore inserted
CC in the membrane, forming cation-selective channels.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: The molecule is folded into four functional domains. Each
CC domain is required for a particular step in the toxicity process.
CC Domain 1 contains two calcium ions and the proteolytic activation
CC site. Cleavage of the PA monomer releases the subdomain 1a, which
CC is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is
CC part of the remaining 63-kDa fragment (PA63) and contains the
CC binding sites for LF and EF. Domain 2 is a beta-barrel core
CC containing a large flexible loop that has been implicated in
CC membrane insertion and pore formation. There is a chymotrypsin
CC cleavage site in this loop that is required for toxicity. Domain 3
CC has a hydrophobic patch thought to be involved in protein-protein
CC interactions. Domain 4 appears to be a separate domain and shows
CC limited contact with the other three domains: it would swing out
CC of the way during membrane insertion. It is required for binding
CC to the receptor; the small loop is involved in receptor
CC recognition.
CC -!- PTM: Proteolytic activation by furin or a furin-like protease
CC cleaves the protein in two parts, PA-20 and PA-63; the latter is
CC the mature protein. The cleavage occurs at the cell surface and
CC probably in the serum of infected animals as well; both native and
CC cleaved PA are able to bind to the cell receptor. The release of
CC PA20 from the remaining receptor-bound PA63 exposes the binding
CC site for EF and LF, and promotes oligomerization and
CC internalization of the protein.
CC -!- MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were
CC performed that showed that the residues present in the small loop
CC of domain 4, and not the ones in the large loop, are involved in
CC receptor recognition.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

```
Query Match 99.4%; Score 3750; DB 1; Length 764;
Best Local Similarity 99.5%; Pred. No. 1.4e-179;
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQNRLLNESESSQGLGYFDLNFQAPMVVTSSTGDLSPSELENIPSENQYF 60
DB 30 EVKQNRLLNESESSQGLGYFDLNFQAPMVVTSSTGDLSPSELENIPSENQYF 89
QY 61 QSAIWGFTKVKKSDYFATFADNHVMTWDDQVINKASNKIRLEKGLYQIKIY 120
DB 90 QSAIWGFTKVKKSDYFATFADNHVMTWDDQVINKASNKIRLEKGLYQIKIY 149
QY 121 QRENTEKGLDFKLYWTDSONKKEYISSDNLQLPKQKSSNRKSTSGPTVPDRN 180
DB 150 QRENTEKGLDFKLYWTDSONKKEYISSDNLQLPKQKSSNRKSTSGPTVPDRN 209
QY 181 DQIPSPLEVEGTVDKKRTPLSPWISNHEKGLTKYKSPKSWTASDPYDFEYV 240
DB 210 DQIPSPLEVEGTVDKKRTPLSPWISNHEKGLTKYKSPKSWTASDPYDFEYV 269
QY 241 GRIDKNVSEARHPIVAAPYPIVHVDENIILSKNEDQSTQNTDSTRISKNTSRHT 300
DB 270 GRIDKNVSEARHPIVAAPYPIVHVDENIILSKNEDQSTQNTDSTRISKNTSRHT 329
QY 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 330 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 389
QY 361 NANIRVNTGTAPIYVNLPTTSLVLCNKQTLATIKADENQLSOILAPNNYPSKNLAPIA 420
DB 390 NANIRVNTGTAPIYVNLPTTSLVLCNKQTLATIKADENQLSOILAPNNYPSKNLAPIA 449
QY 421 LNAQKQDASPTPTMNYNQFLEKTKQLRLDQVYGNVATYFNFGVRVDTGNSWSEV 480
DB 450 LNAQKQDASPTPTMNYNQFLEKTKQLRLDQVYGNVATYFNFGVRVDTGNSWSEV 509
QY 481 LPQIQETARIIFNGKDLNLVRRITAAVNPSPLETTKPDMLTKALKIAPFNENPGLN 540
DB 510 LPQIQETARIIFNGKDLNLVRRITAAVNPSPLETTKPDMLTKALKIAPFNENPGLN 569
QY 541 QYQKQDITEFDNFQDQTSQNIKNQLAELNATNITVLDKIKLNAMNILLRDKRPHYDR 600
DB 570 QYQKQDITEFDNFQDQTSQNIKNQLAELNATNITVLDKIKLNAMNILLRDKRPHYDR 629
QY 601 NNIAVGADESVEKAEHREVSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
DB 630 NNIAVGADESVEKAEHREVSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 689
QY 661 DMLNLSLRQDGKTFIDFKKYNKPLIYISNPNKYVAVTKENTIINPSENGDTSTNG 720
DB 690 DMLNLSLRQDGKTFIDFKKYNKPLIYISNPNKYVAVTKENTIINPSENGDTSTNG 749
QY 721 IKKILFSGKYBIG 735
DB 750 IKKILFSGKYBIG 764

RESULT 2
YPB1_BACAN
ID YPB1_BACAN STANDARD; PRT; 204 AA.
AC P13422; Q9X377;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein pXOI-111.
GN pXOI-111.
OS Bacillus anthracis.
OG Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 6.2%; Score 235.5; DB 1; Length 204;
Best Local Similarity 34.8%; Pred. No. 1.9e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNLIRDKRPHYDRNNIAVGADESVEKAEHREVSSTEGLLNIDKIRKILSGYVIE 646
DB 1 MNLIRDRP-YHYDNGNIVGVDDSYLKNAYKQILNWSDDGVLNLDVQNALSGYMLQI 59
QY 647 EDTE-----GLKEVINDRYDMLNLSLRQDGKTFIDFKKYNKPLIYISNPN 693
DB 60 KPSNHLTNSPVITLAKGDSGVGELYRVL-----DGTGFLDFNKFDEWRSLV-DPG 112
QY 694 YKVNVAATKEN-TIINPSENGDTSTNGIKKILFSGKYBIG 734
DB 113 DDVVVYAVTKEDFNVAVTRDENGNI-NKLKNTLVLSGKIKEI 153

RESULT 3
MSPI_PLAFK
ID MSPI_PLAFK STANDARD; PRT; 1630 AA.
AC P04932;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMW5A) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman M., Bore N., Hyde J.E., Scaife J., Certa U.,
RA Stunnenberg H., Bujard H.;
```

"Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level.";
EMBO J. 4:3823-3829 (1985).
[2]
REVISED, SEQUENCE FROM N.A.
Pan W., Tolle R., Bujard H.,
Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
-!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
(Potential).
-!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.
-!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
EMBL: X03371; CAA27070.1; -
InterPro: IPR006209; EGF-like.
Pfam: PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.
FT DOMAIN 67 84 TRIPETIDE SG(TP) REPEAT.
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;
Query Match 4.9%; Score 184.5; DB 1; Length 1630;
Best Local Similarity 20.5%; Pred. No. 0.11;
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;
QY 5 ENRL-LNESESSSGLLGYFSDLNFCQAPMVVTS*TGDLSTPSSSELENIPSENQVFSQA 63
Db 918 ENLSLGNKNVYQELIGQSSP-NP-----YKILKSDFFYNE 956
QY 64 IWSGFIKXKSDYTFATSNADNHVMTWVDQEVINKASNS-----NKIRLEKGRLY- 114
Db 957 SFTNFVKSADD-----INSLNDSKSKLEEDINKLKTQLSFDLYNKYKLLERLFD 1011
QY 115 -----QIKIQYQRENTKGLDFKLYWTDQSKQKEVSSDNLQLPELKOKSSNSRKKES 168
Db 1012 KKTVGKYKWKIKLLKQLESK---NSLNPKGVLI---QNFVFNKKKEAIEAETEN 1067
QY 169 TSAGTVPDRDNDGIPDSLEVEGYVDVQNKRTFLSPWISNIHKKGLTKY----KSSP-- 223
Db 1068 T-----LENTKILLKHY-----KGLVKYNGESSPLK 1094
QY 224 ---EKWSTASDPYSDPE-----KVTGRIDKNVSPKAR-----HPVAVAYPIVHVD 265
Db 1095 TLESSEITQEDNYSLENFKVLSKLEGLKNDLNLEKKKLSYSSGLHLLHIA-----E 1147
QY 266 MENILSKNEQDSQNTDSETRTISKNTSTSRHTSEVHGNABYHASFDFDGGSVSAGFS 325
Db 1148 LKEVIKKN-----YTCNSPSENNT-----DVNNALSEYKFLPEGTDAIVWS 1191

QY 326 NSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPFYNVLPPT---- 381
Db 1192 ESGDTLSQSQPKPASTHVGAE--NTITTSQ-NVDEVDVVIIVIFGESEEDYDDL 1247
QY 382 -SLVGLKQOTLATIKADENQSLAPNNYPSKNLAPIA---LNAQKDASTPTTMVYN 437
Db 1248 GQVVTGEAVTPSVI---DNILSKI---ENEVLYLKPLAGVYRSLKQLENVWTFVNV 1301
QY 438 -----QFLELEKTKQLRLDQV-YGNIAF-----YNFENGVRVVDTSQNWSEVL 481
Db 1302 VKDILNSRNKENPKV-LESDLPYKDLTSSNVVVDPKYFLNKEKDKFLSSVNYIK 1360
QY 482 PQIQETARIIFNGKDLNLVERRIAAVNPSPDL-----ETTKPDM-TLKEALKIARGF 533
Db 1361 DSID-----TDINFA-----NDVLGYKILSEKYSKSLDSIKYINDKQGE 1401
QY 534 NEFN-----GNLQVQGDITE-----DFNFQDQTSQNTKQNLABELNATNYT 576
Db 1402 NEKYLFLNNIETLYKTVDKIDLFIHLEAKVLNVYTK---SNVEVKIKELN--YLKT 1456
QY 577 VLDKIKLNAMN--ILIRKRFHYDRNN-----IAGV-ADESVVKEAREVINSSTEGLL 628
Db 1457 IQDKLADFKNNFVGIADLSTDYNNHLLTKFLSTGMVFEENLAKTVLSNLLDGNLQGM 1515
QY 629 LNIDKD--IRKILSGYVIEIDTEGLKEVINDRYDMLNLSLRQDQKTFIDPKYNDKLP 686
Db 1516 LNISSQOCVKK-----QCPQNSGGFRHLDE-----REECKILNKYQEGDKC- 1557
QY 687 LYISNPNYKVNVAVTKENTINPSENG-----DTSTNGIKKI 724
Db 1558 --VENFNPTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1590
RESULT 4
MSPI_PLAFW STANDARD; PRT; 1639 AA.
AC P04933;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens) (PMWSA) (P195).
GN MSP-1.
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86014355; PubMed=2995820;
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V., Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.I., Freeman R.R.;
RT "Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum merozoites.";
RL Nature 317:270-273 (1985).
RN [2]
RP REVISIONS.
RA Holder A.A.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential).
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC

DR EMBL; X02919; CAA26676.1; --
DR PIR; A24594; A24594.
DR PIR; S05603; S05603.
DR PDB; 1CEJ; 28-MAY-99.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 1.
DR Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1639
FT CARBOHYD 116 116
FT CARBOHYD 116 116
FT CARBOHYD 268 268
FT CARBOHYD 764 764
FT CARBOHYD 768 768
FT CARBOHYD 783 783
FT CARBOHYD 844 844
FT CARBOHYD 920 920
FT CARBOHYD 964 964
FT CARBOHYD 1058 1058
FT CARBOHYD 1165 1165
FT CARBOHYD 1174 1174
FT CARBOHYD 1445 1445
FT CARBOHYD 1526 1526
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F66 CRC64;

Query Match 4.8%; Score 184.5; DB 1; Length 1639;
Best Local Similarity 20.5%; Pred. No. 0.11; Mismatches 138; Indels 265; Gaps 47;
Matches 170; Conservative 138; Mismatches 138; Indels 265; Gaps 47;

QY 5 ENRI-LNESSSQGLGYFFDLNFOAPMVVTSSTTGLSIPSSSELENIPSNQYFQSA 63
DB ENILSLGKNKIYQELIGKQSSSE-NF-----YEKILKSDTFYNE 965
QY 64 IWSGFIKVKKSDYTFATSDNHTVMVDQEVINKASNS-----NKIRLEKGLY- 114
DB SFTNFVSKADD-----INSLNDESKKXUEEDINKLKTQLSFDLYNKYKILKLERLD 1020
QY 115 -----QKIQYQRENPTKEGLDLYWTSQNKKEVISSDNLQPLKQKSSNSRKRKS 168
DB KKTVGKYKMQIKKLTLLKEQLSKL-NSLNNKPHVL-QNFSVFNNKKKEAIAETEN 1076
QY 169 TSAGPTVPDRNDGIPDSLEVEGVTVVKNKRTFLSPWISNIEHKKGLTKY--KSP-- 223
DB T-----LENTKIILKHY-----KGLVYKNGESSPLK 1103
QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPEAR-----HPLVAAPYVHVD 265
DB TLSESIQTDENYASLENPKVLSKLEGLKDNLNLEKKLSYLSGLHLIA-----E 1156
QY 266 MENIILSKNEPQSTQNTDSETRTISKNTSTSRHTSEVHGAENAVHAFDYGGSVAGFS 325
DB 1157 LKEVIKKNK---YTCNSPSENNT-----DVNNALSYKPKFLPGTDVATVVS 1200
QY 326 KNSNSTVALDHSLSLAGERTWAEITMGLTADTARLNANIRVNTGTAPIYVNLPTT--- 381
DB 1201 EGSOTLQSQSQPKPASTHVGAE-----NTIITSQ-NVDEVDVLIIVIFGESEDDDL 1256
QY 382 -SLVLGKQKQTLATIKADENQLSILAPNNYFSPKXNLAPIA---LNAQKDAASPTITNNYN 437
DB 1257 GQVVTGEAVTPSVI---DNILSKI---ENEVEVLKPLAGVYRSLKKQLENVMTNVN 1310
QY 438 -----QFELEKTKQRLDQV-YGNIAT-----YFNGRVRVDTGSNWSEVL 481
DB 1311 VKDILNSRFNKGKFNKVN-LESDDLIPYKDLTSSNVVVKDPYKFNKEKDFLSYNIYK 1369
QY 482 PQIQETARIIPNGKDLNVERRIAANVNSDPL-----ETTKPDM-TLKEALKIAPGF 533
DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSKLDLSIKKYINDKQGE 1410
QY 534 NFPN-----ONLYQKQKITE-----PFNFQDQTSQNIKNQLAEIENATNIYT 576
DB 1411 NEKILPFLNIIETLYKTVNDKIDLFVHLEAKVLNITYEK-----SNVEVKIKELN-YLKT 1465

QY 577 VLDKIKLNAXN--ILIRDKRFHYDRNN-----IANG-ADESVVKEAHRREVINSSTEGLL 628
DB 1466 IODKLADPKONNFVGIADLSTDYNNHNLTKFLSTGMVFENLAKTFLVSLNLDGNLOG-M 1524
QY 629 LNIKDQ--IRKILSGYIVIEIDTEGLKEVINDRYDMLNLISSLRQDGKTFIDFKKYNKLP 686
DB 1525 LNIHQHCVKK-----QCQNSGCFPHLDE-----RECKLLNYKQBGDKC- 1566
QY 687 LYISNPYKVNYYAVTKNTIINPSENG-----DTSTNGIKKI 724
DB 1567 --VENPNPTCN-----EN-----NGGCDADAKCTEEDSGNSG-KKI 1599

RESULT 5
YME7_YEAST
ID YME7_YEAST STANDARD; PRT; 1658 AA.
AC Q03661; Q04988;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE Hypothetical 187.1 kDa protein in GUL1-ERG8 intergenic region.
GN YMR219W OR YME261.13 OR YMS955.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AS972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels P., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RT Nature 387:90-93(1997).
RL -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49809; CAA89934.1; --
DR EMBL; Z49839; CAA90190.1; --
DR PIR; S55101; S55101.
DR Germline; 142894; --
DR SGD; S0004832; ESCI.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
KW Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;

Query Match 4.5%; Score 169; DB 1; Length 1658;
Best Local Similarity 18.6%; Pred. No. 0.66;
Matches 159; Conservative 140; Mismatches 274; Indels 284; Gaps 41;

QY 1 EVKQNRLLNESGES---SSQGLLYGYFDLNFQAPMVVTSSTTGLSIPSSSELENIPSEN 57
DB 290 EMELEDDIDVESDAEKDESQGAEGIEHS-VDFSKYMPRTDNTKIPVIEKSEDEHKVHQ 348
QY 58 QYFQSAIWS-GFTKVKKSD-----YTPATSDN-----HVTMWVDDQEVINKASNSK 105
DB 349 RYSGDGFDFGVSNIQVDDSEDEESQAESYANAEVYHNEHELDKELIEDIESDS 408
QY 106 IRLEKRLYQIKYQRENPTKEGLDFKL-----YWTDSQKKEV 145
DB 409 -----ESQSAQSESGSEDDPFYKMKNEKSTSEETNTSESDQGFADATKNKV 459
QY 146 ISSDNLQLPEL-----KQKSSNSRKRKRSAGTVPDRNDGIPDSLEVE 190

Db 402 SFLANKESP-----ADIQACVGFALKALIASANAPLDDISIGNSLTRQSLASPESTAKLV 457
 Qy 351 GLNTADTARLNANRYVNTGTAPYVNVLPFTSLVLGKQNTLATIKADENQLSOLAPNNY 410
 Db 458 DI-----MINQRCALNT-----TVSIVI-----ELIRKNSDYDQVNLTTTIXT 498
 Qy 411 YPSKNLAPIALNAOKDASSTPTVNNYQFLEKTKQLRLDQVYGNATYVNFGRVR 470
 Db 499 HPPSNRDPYLYGLLRKFSNHLSDFFQIILDENDANIPLHENQLHKEFKPLGFERFKV- 557
 Qy 471 VDTGSNNEVLPQIQETIARIIFNGKDLNVLVERIAAVNP--SDPLETTKPDVTL--KEA 526
 Db 558 -----VELLAELHCSNGLMNSKRAERARRDKVRQSLSHLQDALNDLSIEKEQ 610
 Qy 527 LKIAFG-----FNEPNGMLQVQKDI-----TEFDNFD-----QOTSQNIKNQLAE 568
 Db 611 LKTHSPTRDTHDLKNNNGKIDNNDNDDSDYGEIDESPFIPIYNNKQIKLRTDP 670
 Qy 569 LNATNIYTVLQKILNAKQNTLIRDKRPHYDRNNIAVGAD-----ESVWKEAREVIN-- 621
 Db 671 TVGT-----YKSNVNIKD-----CFQNNELFLTHPWNFWHNVIFDIQIIFNGR 716
 Qy 622 ---SSTEGLLNID--KDIRKILGYIVIEB---DTEGLKEVINRDYDMLNISS---LR- 669
 Db 717 MDFSYSNSELVLSFLNLSYGFMTDIVISDEKGTDSRFSVIRDPNFDPKITITDFILRG 776
 Qy 670 -QDGKTFIDFKYN-----DKLPLYISN-----PNYKNVY 699
 Db 777 YQDSYKFYELRKQNLGYMGHIVLIAEYVVKSKYKVDYISRYSNLQTEBWQYISEVL 836
 Qy 700 AVTK-----ENTINPSENGDTSTN 719
 Db 837 NETRMYSKILGGSYIDDGNGNLIPLQPDNTTVLTP--NGDASNN 880

RESULT 7

BDX_CLOBO
 ID - BDX_CLOBO STANDARD; PRT; 1276 AA.
 AC P19321;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)
 DE (Bontoxilysin D)
 GN BCTD.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BVD/-3;
 RX MEDLINE=91016853; PubMed=2216736;
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
 RA Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;
 RT "Nucleotide sequence of the gene encoding Clostridium botulinum
 RT neurotoxin type D";
 RL Nucleic Acids Res. 18:5556-5556 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CB16;
 RX MEDLINE=93042276; PubMed=1420572;
 RA Sunagawa H., Ohnaya T., Watanabe T., Inoue K.;
 RT "The complete amino acid sequence of the Clostridium botulinum type D
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RT phase d-16 phi genome.";
 RL J. Vet. Med. Sci. 54:905-913 (1992).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC STRAIN=D-SA, and D-1873;
 RX MEDLINE=89339741; PubMed=2668193;
 RA Morishishi K., Syuto B., Kubo S., Oguma K.;

"Molecular diversity of neurotoxins from Clostridium botulinum type D strains."; Infect. Immun. 57:2886-2891 (1989).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94230352; PubMed=8175689;
 RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
 RA Roques B., Fyke E.M., Suedhof T.C., Jahn R., Niemann H.;
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
 RT F botulinum neurotoxins and tetanus toxin.";
 RL J. Biol. Chem. 269:12764-12772 (1994).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS--LEU-61 BOND OF
 CC SYNAPTOSOMAL VESICLES AND -2.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H). The light chain has the pharmacological activity,
 CC while the N- and C-terminal of the heavy chain mediate channel
 CC formation and toxin binding, respectively.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin; Types A, B, C1, D, E, F, and G.
 CC -!- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 CC BACTERIOPHAGE.
 CC -!- SIMILARITY: Belongs to peptidase family M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X54254; CAA38175.1; -
 CC EMBL; S49407; AAB24244.1; -
 CC PIR; S11455; S11455.
 CC HSP; P10845; 3BTA.
 CC MEROPS; M27.002; -
 CC InterPro; IPR008985; ConA like lec gl.
 CC InterPro; IPR002160; Kunitz legume.
 CC InterPro; IPR006025; Pept_M_Zn_BS.
 CC InterPro; IPR000395; Peptidase_M27.
 CC Pfam; PF01742; Peptidase_M27; 1.
 CC PRINTS; PR00760; BONTOLYLISIN.
 CC PRODOM; PD001963; Bontoxilysin; 1.
 CC PROSITE; PS00142; ZINC PROTEASE; 1.
 CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 CC CHAIN 1 442 BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.
 CC CHAIN 443 1276 BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.
 CC METAL 229 230 ZINC (CATALYTIC) (BY SIMILARITY).
 CC ACT_SITE 230 233 ZINC (CATALYTIC) (BY SIMILARITY).
 CC METAL 437 450 INTERCHAIN (PROBABLE).
 CC DISULFID 15 16 ND -> PV (IN STRAIN D-SA).
 CC VARIANT 17 18 ND -> LQ (IN STRAIN D-1873).
 CC VARIANT 452 452 K -> Q (IN STRAIN D-SA).
 CC VARIANT 457 457 R -> T (IN STRAIN D-SA).
 CC VARIANT 457 457 R -> F (IN STRAIN D-1873).
 CC VARIANT 462 462 A -> D (IN STRAIN D-1873).
 CC VARIANT 489 489 K -> N (IN STRAIN CB16).
 CC VARIANT 489 489 N -> K (IN STRAIN CB16).
 CC VARIANT 644 644 Q -> R (IN STRAIN CB16).
 CC VARIANT 1122 1122
 CC SEQUENCE 1276 AA; 146871 MW; C1EC50F46C8233E2 CRC64;

Query Match 4.4%; Score 166; DB 1; Length 1276;
Best Local Similarity 20.7%; Pred. No. 0.65;
Matches 165; Conservative 117; Mismatches 290; Indels 224; Gaps 41;

QY 44 SIPSELENIPENOYFOSATWSGFIKVKSDVEYFATSADNHTVMVDDQEVINKASNS 103
DB 310 TIPSWISNI---DKY--KKIFSEKYNFKONTGNFVNIDKENSLSDLTNWSEVYVS 364
QY 104 NKIRLEK-----GRLYQIKQVQRENPIEKGDLFKLYWTDSONKKEV 145
DB 365 SQYVKNRTHYPSRHYLFFVFNILDDNIYTRDGF---NLTKGNFIE----- 409
QY 146 ISSDNLQL-PEUKOKSSNRKRSAGTVPDRDNDGIPDSLEVEGYVDVKNKRTFLS 204
DB 410 NSGQNIERNPALOKLSSESVDLFTKVLRLTKNSRD-----DSTCIKVKNNRL--- 458
QY 205 PWISN-----IHKKGKLTGYKSSPEKMSSTASDPYS-DFEKVGRIDKNVSPPEARPLV 256
DB 459 PYVADKDSISQIFENKIITD-ETVQNY---SDKFSLDESILDCQVP--INPEVDPL- 511
QY 257 AAYPIVHVDMENILSKNE---DQSTQNTD-----SETRTISKNTS--TSRTHUSEV 303
DB 512 ----LNVNMEPLNIPGEIIVFYDITKYVDYLSYVYLESQKLSNNVENLTITSVEEA 567
QY 304 HGNAEVHSAFF-----DIGGSVSAFSSNSSTVAIDHSLSLAGERTWAETM----- 350
DB 568 LGVSNKIYTFPLSLAEKVNKGQVAGLFLNANEVVEFTNIMKDTLTKISDVSIIPIY 627
QY 351 ---GLNTADTA-RLNANIRYVNTG-----TAPIYVLPPTSJVLGKNQTLATIK 395
DB 628 IGPALNIGNSALRGNFQAFAGVAFLEGPEFTIPALGVFTYSSIQEREKIITIE 687
QY 396 -----ADENQLSQILAPNNY-----YPSKLNAPIALNAQKDSSTPTM 434
DB 688 NCLEQRVKWQDSYQWVSNLRSRTTQFNHINYQWYDS-----LSYQADAIAKAKIDL 740
QY 435 NYNQFLEK-----TQQLRLDTPQVQNI-----ATYFENGVRVDTGSN 476
DB 741 EYKYSQSKENIKSOVENKNSLDVKISEAMNINKFIRECSVTYLFKN----- 790
QY 477 WSEVLQIQETTARIIFNGKD--LNL-----VERRIAAVNPDPLETTKPDMTL 523
DB 791 ---MLPKVIDELNKFDRTKTELINLIDSHNIIIVGEVDRKAKYNES--FENTWP---- 841
QY 524 KEALKIATFG-FNEPKNLQVQKDIETEPFNFQOTS-----QNIKNQALNATNI-YTV 577
DB 842 -----FNIFSYTNLSL---KDIINEYFNSINDSKILLSQNKALVDTSGYNAEVRV 891
QY 578 LDKIKLNAXONILIRKFRHYDRNNIAGVADSQVVKAEHREVINGSTEGLLNIDKIRK 637
DB 892 GDNVQLN---TIYTNDFKLSSGDKIIVLNNNIIYSA---IYESSVSFWIKISKDLTN 945
QY 638 ILSGYIV-BIEDEGLKEVINDRYDMLNIISSLRQDGKTFIDFKYNDKLPY----ISNP 692
DB 946 SHNEYTIINSIEQNSGWKLCIRNG---NIEWILQDVN-----RKYKSLIFDYSLSGHT 996
QY 693 NYKVNVYAVKENTII 708
DB 997 GYTNKWFVITNNIM 1012

RESULT 8

YJL3 YEAST
ID YJL3 YEAST STANDARD; PRT; 1803 AA.
AC P47024; P87192;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transposon Ty4 207.7 kDa hypothetical protein.
GN Ty4B OR YJL113W OR J0780.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97103775; PubMed=8948101;
RA Czepluch C., Kordes E., Pujo A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,
RT STP10, GCD14, RPE1, FHO86, NCA3, ASF1, COT7, GZF3, two tRNA genes,
RT three remnant delta elements and a Ty4 transposon.";
RL Yeast 12:1471-1474 (1996).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; Z49389; CAA9409.1; -.
DR SGD; S0003649; YJL113W.
DR InterPro; IPR001594; Rve.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00665; rve; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR Transposable element; Hypothetical protein.
SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;

Query Match 4.4%; Score 165; DB 1; Length 1803;
Best Local Similarity 19.5%; Pred. No. 1.2; Mismatches 274; Indels 270; Gaps 41;
Matches 162; Conservative 123;

QY 1 EVKQNRLLNESSSQGLLYFSDLNFPQAPMVTSTTGDLSPSSLENIPISENQYF 60
DB 1078 QLAKTN---HETSPKESGIG---TNVGFANTNNEISLTKGTSLPTIKLESINHH--- 1128
QY 61 QSAINSGFIKVKSDVEYFATSADNHTVMVDDQEVINKASNKIRLEKRLYQIKCY 120
DB 1129 -----SNDY-----STNKVE-----KF 1140
QY 121 QREN---PTEKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTYP 176
DB 1141 EKENHPPPIEDIVDM-----SDQTMESNCODGNLKLKELKVTDKVPTDNGTNSPR-- 1193
QY 177 DRNDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKK-----GLTKY----- 219
DB 1194 -----LEQNEASGSPQTVNKSFAFLNKEFSLNKKRKRKHDKNSLTSYLELKKR 1247
QY 220 -----KSSPEKWSADSPYSDPEKVTGRIDKNVSPPEARPLVAAYPVHVDMENILSK 273
DB 1248 SKNRVKLIIPDNMETVSAPKIRAIYNEAISKNPDLKERHEYKQAY---HKLQNLKDMK 1304
QY 274 NEDQSTQNTQSETRTISKN---TSTSRTHSEVHGNABHASFFDIGSVSAGFSNSST 331
DB 1305 VFVDVVKYRSRSE---IPDNLIVPTNTIFTKRRNG---IYKARIVCRGDQSTPDYSVIT 1358
QY 332 VAIDHS-----LSLAGERT-WAETMGLNTADT-ARLNANIRYVNTGTAPIYVLP-TTSL 383
DB 1359 ESLNENHIKFLMIANNRMFMKTLIDINHAFLAKLEE-----IYIHPHDDRC 1408
QY 384 VLGNQTLATIKADENQ-----LSQI-LAPNNYYP-----SKNLAPIALNAQKAS 428
DB 1409 VVKLNKALYGLQSPKEWNDHLRQYLNIGLKONSYPGLYQOTEDKNLM-IAVYVDDCVI 1467
QY 429 STPTIMYNOFL-ELEKTKLR-----LDTDOVYGNIAIYNFENGVRV----- 469
DB 1468 AASNEQRLDEFINKKFNENFKLTGTGLIDVLDTD-ILGMDLYNKKELGTIDITLKSFIN 1526
QY 470 RVDTCGNSWSEVLPQIQTETARIIFNGKDLNLFERRIAAVNP-SDPLETTKPD-----MTL 523
DB 1527 RMD--KKYNEELKIKRKSIPHMSTYK-----IDPKKQVLQMSSEERQGVKL 1573


```

QY 524 KEALKIAGFNBNGLQYQGGDIETFFDFNFQQTSONIKNQLAELNATNIYTVLDKIKL 583
Db 1574 QQLL-----GELNVV-RHKCRDYIEFAVKVAVLVNYPHERVFMVYKIIQYL-- 1620
QY 584 NAKMILLIDKEPHFDN-----NIAVGADSVVKEAHE-----VI-----NS 622
Db 1621 -----VRYKDIGHYDNCNKKVIAITDASVGEYDAQSRIGVILWYGMNIFNVYGNK 1675
QY 623 STEGLNLNDKIDKILSGY-----IVEIEDT-----EGLK----- 653
Db 1676 STNRCVSTAEALHAIEYADSETLKVTLKELGEGNDNDIWIITDSPAIGLNRSYQQ 1735
QY 654 -----EVINDRYDMNLNSSLRQDGKTFI-----DPKKY 681
Db 1736 PKXKFTWIKTEIKKIKESIKLKITGKNIGADLLTKPVSASFKR 1784

RESULT 9
ID SPOF SCHPO STANDARD; PRT; 1957 AA.
AC Q10411; Q9USE9;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sporulation-specific protein 15.
GN SPO15 OR SPAC1F3.06C
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
RX MEDLINE=20107136; PubMed=10639340;
RA Ikemoto S., Nakamura T., Kubo M., Shimoda C.;
RT "S. pombe sporulation-specific coiled-coil protein Spo15p is localized
RT to the spindle pole body and essential for its modification."
RL J. Cell Sci. 113:545-554(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hanlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Boerzj M., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
RN [3]
RP SEQUENCE OF 705-871 FROM N.A.
RC STRAIN=968 h90;
RX MEDLINE=20223868; PubMed=10759889;
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Harauchi T.,

```

```

RA Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
RT fission yeast cells by the use of a GFP-fusion genomic DNA library."
RL Genes Cells 5:169-190(2000).
CC -!- FUNCTION: Has a role in the initiation of spore membrane
CC formation.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Spindle pole body.
CC -!- SIMILARITY: Belongs to the MPC70 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z70690; CAA94624.1; -.
CC EMBL; AB027811; BAA87115.1; -.
CC FIR; T38077; T38077.
CC GeneDB SPombe; SPAC1F3.06c; -.
CC Sporulation; Coiled coil.
CC DOMAIN 199 785 COILED COIL (POTENTIAL).
CC FT DOMAIN 804 1235 COILED COIL (POTENTIAL).
CC FT DOMAIN 1320 1471 COILED COIL (POTENTIAL).
CC FT DOMAIN 1481 1723 COILED COIL (POTENTIAL).
CC SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;
CC
CC Query Match 4.3%; Score 164; DB 1; Length 1957;
CC Best Local Similarity 20.4%; Pred. No. 1.5;
CC Matches 177; Conservative 132; Mismatches 328; Indels 232; Gaps 40;
QY 2 VKQENRLINSE-----SSSQGLLGYFPDLNFQ---APMVTSSITGDLSPSSR 49
Db 129 VTQKSNLNLKQVRSKLALEHNGILSLQSSKKQNTSVTLTSEEDVSFQKK 188
QY 50 LENIPSENQYFQSAIN--GFIKV-----KKSDYTFATSDNHNVTWVDDDEVINKAS 101
Db 189 LTNMESNFSKQSEAYDLRSQLLTVTEKLDKEDYKEDVSSIKASLAERQASNKSL 248
QY 102 NSNKIRLEK---GELYQIKYQRENTE---KGLDFKYWTDSQNKKEVISSDNLQPE 155
Db 249 RGEQERLEKLLVSNKTVTLRTQENSLRAECKTLQEKL-----EKCAINEDSKLLEE 302
QY 156 LKQKSSN-----SRKRSSTAGTVPDRDNDGIPDSLEVEGYTVDVKNR----- 200
Db 303 LKENVANYSDAIVHKDKLIEDLSRISEFN-----LKSERDTLSIKNEKLEKLLRNTI 356
QY 201 -TFLSPMISNIHEKKGLTKYKSS-----PEKWTASDPYSDFEK-----VTGRID----- 244
Db 357 GSLKDSRTSNLSQLEENVELKESNRTIHSQLTDAESKLSFQENKSLKGSIDYQNNLS 416
QY 245 -----KNVS---PEARHPLVAA-----YPIVHVDN----- 266
Db 417 SKDRNVQVSQLEEARSSLAHATGKLAENISERDFQNKIKDFEKEIQDLRACLNSSN 476
QY 267 ---ENILSKNEDQSTQNTD---SETRTISKNITSTS-----RHTSTSEVHGAENVHASFDD 315
Db 477 ELKEKSLIDKKQDELNNLREQIKQEKQKVSSESTQSSLSQSLQRLDILNEKKKHEVYESQLNE 536
QY 316 IGGSVSAGFNSNS-----STVAIDHSLSLAGERTWAETMGLTADTARLNANIRYVNTG 370
Db 537 LKGLQTEINSEHLSQLSTLAAKKAATVATNELSESK--NSLQTL-CNA----- 585
QY 371 TAPIYNYLPTTSLVLGKNQ---TLATIKADENQLSQILAPNNYPSKNLAPIALNAQKD 426
Db 586 ---FOEKLAQSVMLQKENEQNFSSLDTSFKKLNESHQELNNHQITIKQL-----KD 634
QY 427 ASSTPITMYNQFLEKTKQLRLDQVYGNATYFNFGRYVDTGNSWSEVLPOIQE 486
Db 635 TSS-----KLQQLQLERANFQKSESTLSDENNDRK-----LLKDEE 672

```

```

QY 487 TTARIIFNGKDLNLVARRIAAVNPSPLETTKPDMLTKEALKI-----AFGNEPENGNIQ 541
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 673 SNKSLIKKQSDVDSLEKNI-----QTLKEDLRKSEALRFSKLEAKNLREVIDLNL- 722
QY 542 YQKQDIT-EFDNFQDOTSONIKQVLAELNATNIYTVLQIKLNKAKWNLIRDKRPHYDR 600
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 723 -KQKHTLEAQRDLHSSUSDAKNTAILSELTKSSSEDKVRLTANVELTQDSK----- 776
QY 601 NNIAGADESVVKEAREVINSSTEGALLNIDKDIR-----KILSGYIVETI-----DT 649
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 777 -----AMQSFSTSLVNSVQS--ISNLYHELADDDHVNQSQNTILLESKLTDC 824
QY 650 EQLKE---VINDRYDMLNTSLRQDQKTIIDFKYNDKLPVLSINPNYKVNYVYKENT 706
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 825 ENITQQNTLIDNVQKLMKHVNQESKV-SELKEVNGKLSLDLKNRLSNLV-AISDNDQ 882
QY 707 IINP-----SENGDT-----STNGIKKI 724
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 883 ILTQLAELSKNYDSLQESQALNGLKSL 911

RESULT 10
FAT1 SCHPO
ID FAT1 SCHPO STANDARD; PRT; 1385 AA.
AC O13735; Q9UNJ0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 18-OCT-2001 (Rel. 40, Last sequence update)
DE Actin interacting protein 3 homolog.
GN FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Munnell K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkert G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vansteenkiste E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carriutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
RN [2]
RP IDENTIFICATION, AND GENE NAME.
RX MEDLINE=20143585; PubMed=10675021;
RA Jin H., Amberg D.C.;
RA "The secretory pathway mediates localization of the cell polarity
RT regulator Alp3p/Bud3p.";
RL Mol. Biol. Cell 11:647-661(2000).
CC -!- FUNCTION: Involved in the organization and/or function of the

```

```

CC actin cytoskeleton.
CC -!- SIMILARITY: TO YEAST BUD6.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z97208; CAB10112.1; -.
DR EMBL; AL109770; CAB52420.1; -.
DR GeneDB SPombe; SPAC15A10.16; -.
DR GO; GO:0005884; C:actin filament; ISS.
DR GO; GO:0016043; P:cell organization and biogenesis; ISS.
DR InterPro; IPR005613; AIP3.
DR Pfam; PF03915; AIP3; 1.
KW Coiled coil; Cytoskeleton.
FT DOMAIN 1009 1096 COILED COIL (POTENTIAL).
SQ SEQUENCE 1385 AA; 154325 MW; ABB3D40C4FF7537 CRC64;

Query Match 4.3%; Score 162; DB 1; Length 1385;
Best Local Similarity 19.2%; Pred. No. 1.2;
Matches 147; Conservative 125; Mismatches 293; Indels 202; Gaps 33;

QY 26 DLNQAPMWTSTSTGDL-----SIPSELENIPISENQYFQSAIWS-----GFIKVKKS 74
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 326 DRKFSPLKTRTPSLTKSLDFGTPSLKSPSLRKSPS--SFVQKDVYSRNSLRISQANRS 383
QY 75 DEYFATFADNHNVTMVDQEVINKAS-----NSN 104
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 384 NVFPGATD---VTRSVSDHRLSSSTINDGEVAPLPQSRSTISSPNSPLSATVPSST 439
QY 105 KIRLEKGRLYOIKIYOQREN-----PTEKGLDFKL-----YWTDSQKK 143
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 440 PILLGRGRSTLSVNVKQFNADGSLNSPNSIRETEEYAAAPKLEDDIADDEVTDATSOR 499
QY 144 EVI-----SSDNLQLPELKOKSS--NSRKKRSTSGAPVVPDRDNGI-----PDSL 187
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 500 ELLERQTKAESSEDTSEIESLQKSLPQVSTQEQEIQSSSVPEASNEIAEKPAVT 559
QY 188 EVEGYTVVDVKNKRTFLSPWTSNIHEKGLTKYKSSPEKWTASDPYDFPKVTGRDKNV 247
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 560 AIESIT-----ERKEEAPVISSEKTESGTSIGTSDTKGGLANFENDSLSELERLIQQN- 613
QY 248 SPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTQSETRTISKTSTSTRTHTSEVHNA 307
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 614 NAEQDEPSYKPKYEYSER-----SGSEDEFKSEKDTKGVYINDSDSTQVEESEDKSTP 669
QY 308 EVHASFFDGGSVSAGFSNNSSTVAID-----HSLSLAGERTWAETMGLNTADTARLNAN 363
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 670 NTGA-----SAKLNDPSSITITVSDVYKPKASPVEITEPPSSALVSATSPTTNVP 720
QY 364 I--RYVNTGTAPVNVLPVTTSLVLGKQNTLATIKADENQLSOLILAPNNYPSKNLAPIAL 421
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 721 IVPEAVHLSTA--FSTAPVSTIV-----SNISPLPTVAP--P 753
QY 422 NAQKQDASSTPITM--NTNQFLELEKT--KQLRLDTQVYGNIAIYINPENGVRVYDTGSNW 477
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 754 NVSGSPSETPISEKPEKVPVVSQTEKALPKPLGVDTKEY--FLRYNNQTKKVKVESPLSN 811
QY 478 SEVLPOQETTARIIFNGK--DLNVERRIAANVPDPLETTKPDMT-LKEALKIAPGFN 534
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 812 ANELGELFSNVYKISFSGDSYELNIED-----PDTKISYLLLEDLSLKSLVSPMFK 864
QY 535 EPNGNLQ---YQKDIITEFDFNFQDOTSONIKQVLAELNATNIYT-VLDIKLNKAKNII 590
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 865 EQDANKKREDPFSGEVSAIOHSAQNTLDDHVNTTTHESPSSAFTLEELKA----- 917
QY 591 IRDKRFHYDRNNITAVGADESVVKEAREVINSSTEGALLNIDKDIRKILSGYIV-EIEDT 649
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 918 -----IEQNIST-----NHTNDSAAKSSDSDHSHKANNFVSPPSIDH 954

```


CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34956; AAA25423.1; .
DR PIR; JQ0894; JQ0894. ABC transporter.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR003439; GTP-binding_dom.
DR InterPro; IPR003405; SMC C.
DR InterPro; IPR003395; SMC N.
DR Pfam; PF02483; SMC C; 1.
DR Pfam; PF02463; SMC N; 1.
DR TIGRfams; TIGR00650; MG442; 2.
KW ATP-binding; Coiled coil. ATP (POTENTIAL).
FT NP_BIND 32 39 COILED COIL (POTENTIAL).
FT DOMAIN 169 224 COILED COIL (POTENTIAL).
FT DOMAIN 231 400 COILED COIL (POTENTIAL).
FT DOMAIN 569 821 COILED COIL (POTENTIAL).
FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).
SQ SEQUENCE 979 AA; 110566 MW; 30D51C56B56280F4 CRC64;

Query Match 4.3%; Score 160.5; DB 1; Length 979;
Best Local Similarity 19.0%; Pred. No. 0.86;
Matches 176; Conservative 132; Mismatches 342; Indels 275; Gaps 40;

Qy 13 ESSSQGLGYPFSDLNFOAPMVV-----TSSTTGDLSIPSELENIPSE 56
Db 51 EQSAQLRLGLNDDVIFAGSKTVKQKAMVXLTFCNEDAIBETKQIFTSLLKRGQGT 110

Qy 57 NQYFQS-----AIWSGFYKXSDYTTATSDADNHTVMVDDOE--VINKASN 102
Db 111 NEFYNDQPVRYNDIKNLAVESG---ISKSLAISQGTISEIABATPEQRKAVIEEAG 167

Qy 103 SNKIRLEKRLQIKIQYQRENPTKGLDFKLYWTDSQNKKEVSSDNLQLPELKQKSN 162
Db 168 TSKYKLDKEEAQKLR-----TNDAD-ILQGAIKELERQVNSLDK-QASKAKIYLEK 219

Qy 163 SRKRSTAGTPVDRD-----NDGIPDSLEVEGYTVDK-NKRTFLSPWISNHEKKG 215
Db 220 SKALESVEGLIYNDLNFNEKLNLTSLLEVEQQRNDLELNIQYESSISQTVH---- 275

Qy 216 LTKYKSSPEKMWSTADPYSDPEKVTGRID--KNVSP-----EARPLVAAPVHV 264
Db 276 ---FKTEV-----SSIQEITSKLDNLKNAISEINLQEARIEERKLIISGEIV-V 322

Qy 265 DMENII--LSKN-EDOSTQNTDSEITISKNSTSTSTHT----- 300
Db 323 DQKTKIEIKKQVESLKIQINASKQREIELDQOLTRLNKANSKLQENDINKEIGVLL 382

Qy 301 --SEVHGNAEVHASFFDYGVSAGFSNSSTVAIDHSLSLAGERTWAB---TWGLNTA 355
Db 383 KKSAAAANINILKQOFENKSFSGKTKIDNSFLPDGVIGLASELKFVSESPSLAIEIV 442

Qy 356 DTARLN-----ANIRYVNTGTAPIYVNLPTSL-----VLG 386
Db 443 LGAALNQIVMKTSEDVQLQALDFLKNLSGKATFIPLTSIKEREVRDHLVLKQKGFGL 502

Qy 387 KNOTLATIKADENO-----LSQILAPNNYPSKVLAPI-----ALNAQKDSSTPT 433
Db 503 VAKELIEFDQFNKLFGLIGNILVYVNDVNDNANRIAKILDHKYTVLSLGGDLFRPGGT 562

Qy 434 -----MNYNQFLELEKTKQLRLDQ-----VYGNATYFNENGRVVD 473
Db 563 GGSGLERTSILNYDIKIK-EHTNLTAFEDQIHLKIKQOTIYNEIETVNSTIQQVKIEA 621

Qy 474 GS-----NWSEVLQIQE--TTARILNGKDLNVERIAAV-----N 509
Db 622 NSINSKLINILNBNLNLKLNASEIFKEQOQDQSLNLSFDSEKLN-IEKQISLTTLT 680

Qy 510 PSDDL-----ETTKPDMTLK-----EALKIAGFNENGNLQYQGD 546
Db 681 KQRLTNLISEQKGETTKQELDAKRLKUNTOHSDSITEQNAKFLVEQNKELSEHYKL 740

Qy 547 ITE-----PDFNFDOOTS--QNIKNQLAELNATNIYVLDKIKLNKAKNILIRKRFH 597

Db 472 LSESLDTRKELWRKEQLQTVLETLLSDVNQNR-----NVNMTSRSLANGIINVK 525
Qy 335 DHSLSLAGE-----RTWAETWG-----LNTADTARLNANIRYVNTG 370
Db 526 TEKIKISPEVFGUGELIKNDKYKTCAEVIGNSLFIHVVDTEETATLIMNELYRMKG 585
Qy 371 TAPIYVNLPTSLVLGKNGQTIATIKADENQSQILAPNNYPSKXNLAIAQAQDASST 430
Db 586 GRVTF--IPLNRLSLDSVKFSPSTTTQIQTPLIKIKIYEPFEKA-----VKHVP 638
Qy 431 PITVYNNQFLEKTKQLR--LDQDV-YGNAT--YNEFNGRVVDTSNKSSEVLPOI 484
Db 639 IVVKDLGGGLAKKXKHNATLDQDRADKRGVLTGGYLDQKTRLESKLNLSRSQH 698
Qy 485 QETTARIIFNGKDLNLVERRIAAVNPS-----DPLETKPD-WTLK 524
Db 699 KKILBELDFVRNELNIDTKIDQVNGNIRKVSNDRESVLTIIEVVRTSLNKKNEKILE 758
Qy 525 EALK-IAGFNEPENGNOYQKIDTEPDFN-----FDQOTSQNIKNOLA-----E 568
Db 759 ESLNAILIKLEKLNTRTFAOEKLNFTFENDLLQDFDSELSKEERLESUTKEISAHNK 818
Qy 569 LNATN-----IYVLDKIKLNKAMNILIRKRFHYDRNNIAG-----A 607
Db 819 LNITSDALEGITTID--SLNAELSKLIPQENDLESKMSEVGDAFTGLQDELKEIQLE 876
Qy 608 DESVKEAHREVINGST-----EGLL-----NIDKDIRKI 638
Db 877 KESVEKOHENAVLELGTVQREIESIABETNNKLEKANNQORLLKLDNFQKSEKT 936
Qy 639 LSGYIVETEDTGLKE-----VINDRYDM-----LN-----ISSLRQDGK 673
Db 937 MIKKTIVTRRELQORIEIGLLPEDALVNDFSITSDQLLORLNDMNTVIGLKNVK 996
Qy 674 -TFIDPKKNYK-----LPLYSNPNYKVNVAVTENTINPSENG 714
Db 997 RAFENFKFENRRKOLAEASELDESQSDLIIVKLKQKYNV-----DSTFQKSENF 1052
Qy 715 DT-----STNGIKKILIFSK 729
Db 1053 EAVFERLVRPGTAKLIIRK 1072

RESULT 12
P115 MYCHR
ID P115 MYCHR STANDARD; PRT; 979 AA.
AC P41506;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE P115 protein.
OS Mycoplasma hyorhinis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2100;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91138990; PubMed=1825306;
RA Notaricola S.M., McIntosh M.A., Wise K.S.;
RT "A Mycoplasma hyorhinis protein with sequence similarities to
RT nucleotide-binding enzymes.";
RL Gene 97-77-85 (1991).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- DOMAIN: Consists of two putative central coiled-coil regions
CC flanked by putative globular regions at the N- and C-termini.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYCOPLASMA
CC P115.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

741 TLEAASEQSLDLIEQARHFVDSLKKEKELGNVNLAEITEFENVQR-----YOEKKQY 796
598 YDRNNAIVAGDESUVVKEAHRVINSSTEGLLNLDKD-----IRKILSGYVIEDTEGL 652
797 IEELTTAKSKIEEASLDLKKIINKITE--IVNVNHEFNVFVQMPGGGKAIHFT--- 851
653 KEVINDRYDMLNIS---SLRODGRFTIDFKYND-----KPLVYI-- 689
852 -----DKNDILNSGVEISAQPPGKTIKNLRLFSGGERKAIISLLFAILKARPIPLCILD 906
590 -----SNPNYKUNVYAVTKENT 706
907 EVEAALDESINVRYVEFLKLLKENT 931

RESULT 13
ID MDS3_YEAST STANDARD; PRT; 1487 AA.
AC PS3094;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MDS3 protein (MCK1 dosage suppressor 3).
GN MDS3 OR YGL197W OR G1307.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FV1679;
RX MEDLINE=971197971; PubMed=9046087;
RA Coglievina M., Klima R., Bertani I., Delneri D., Zaccaria P.,
RA Bruschi C.V.;
RT "Sequencing of a 40.5 kb fragment located on the left arm of
RT chromosome VII from Saccharomyces cerevisiae.";
RL Yeast 13:55-64(1997).
RN [2]
RP CHARACTERIZATION.
RA Li M.B., Neigeboorn L.;
RL Unpublished observations (XXX-1997).
CC -!- FUNCTION: NOT KNOWN; NEGATIVE REGULATOR OF EARLY MEIOTIC GENE
CC EXPRESSION.
CC -!- SIMILARITY: Contains 3 Kelch repeats.
CC -!- SIMILARITY: TO YEAST YER132C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X91837; CAA62947.1; ..
CC EMBL; Z72719; CAA36909.1; ..
CC PIR; S62048; S62048.
CC GeneOnline; 141245; ..
CC TRANSFAC; T03484; ..
CC SGD; S0003155; MDS3.
CC GO; GO:0005737; Cytoplasm; IDA.
CC GO; GO:0030435; Sporulation; IGI.
CC InterPro; IPR006652; Kelch_rep.
CC Pfam; PF01344; Kelch; 2.
CC KEGG repeat; Repeat; Meiosis.
CC REPEAT 171 226 KEGG 1.
CC REPEAT 234 287 KEGG 2.
CC REPEAT 371 419 KEGG 3.
CC SEQUENCE 1487 AA; 167073 MW; 768AEFBAB796E447 CRC64;

Query Match 4.3%; Score 160.5; DB 1; Length 1487;
Best Local Similarity 19.6%; Pred. No. 1.5;
```



```
QY 55 SENQYFQSAIWGFIK--VKKSDYTPATSDNHVTM----- 89
Db 329 TTNYSASSQKWDLAKAYLYLTKGYD-KTTGTNYAFMKFNLLKPIQNMVTVKATLKYVA 387
QY 90 -----WYDDOEVINK-----ASNSKIRLEKGR--LYQIKIQVOR 122
Db 388 HSYVGTGKATGLMLD---TVNSYDNDNAKVNTWTKPASKNIGKADVHKQWASDYDTAAVKS 444
QY 123 ENPTEKGLDFKLYWTDSONK---KEVSSDNLQLPELKQKSNRKRKSTAGTVP--- 176
Db 445 WNSGGANYGFKLH-TNGNGKEYWKKLIS-----SANSANKPIYEVITPIKGN 491
QY 177 -----DRDNDGIPD---SLE-VEGYTVDVKNKRTFLSPFWISNHEKGLTKYKSSP 223
Db 492 TPTIKAYHNGDSNGYFDISMKKVEGAGYKWIYNGKEYQALSAGNV-----TSWTKG 545
QY 224 EK-WSTASDPYDFEKVGTGRID-KNVSPEARPLVAAYPIVHVDMENILSKNE----- 275
Db 546 KKIWTSIAEIAKRYKL--HLGDKGAEALALDP---SPVYKNSGGSYATSKNTWIGVSA 599
QY 276 --DQSTQNTDSETRITISKNTSTSRTHSEVHGNAEVAHSPFDIGGSVAG-----FS 325
Db 600 IFDQGGGASAPAKPVPVWVQAQAPSAKGYNGNA-TGYFDLSWKA VSGATGYKVQVFN 658
QY 326 NNSNSTVAIDH--SLSLACERTWAETMGLNTADTA-----RLNANIRVYNTGTA 372
Db 659 KGKFETLDLGNQTSWTKKLIWPTSAAETKAGYALHLKDGSGAELPINPGTYKNAGGD 718
QY 373 PIYNVLFTSLVGLKQNTLTIKADENQSLQILAPNN-----YVPSK-----NLA--- 417
Db 719 GAKRNSYFKIIAYNKDGEALASPAATPALPDIAEPKNVTGYLYNTKSSQGYVNLWEK 778
QY 418 -----PIALNAQDASSTPI-----TMYNQFLELEKTK-----QJRL 450
Db 779 VQNAKYKYNVINGKEYQSFVGDGDADHWTQNKNIWPTSEIKAGSYKHLTDGKGGLAL 838
QY 451 DTDQVYGN-----IATYN--FENGREVDVTSKNWSEV-L 481
Db 839 DPSPVYNNANGYKGNYSFTLVAYDANGETIPTAPENPTFHEGAEFLGTEEYWSIID 898
QY 482 P-OIQTETARIIFNGKDLNL-----VERIAAVNPSPDL-----ETT----- 517
Db 899 PSGLNGATGNVINEEDLSDIGRFGGLSTYNSLSDHFLFQGGWYDAETSVID 958
QY 518 -----KPDWTLKEAL-KIAG-FNEPENGMLQYQGDITEFDNFQDQTSQNIKQLAEL- 569
Db 959 GGAMY-DEDATTHRFPKADGTYPPTGYVLELTETADQFILKTDQINAYFNKKGKLQ 1018
QY 570 -----NATNIVTLDKIKLNAKMILLIRDKFHHYDRN-----NIAVGADES 611
Db 1019 KVVGDHNNATVYINDKNQLTAITDASGRKLFTFYDENGHVTSTGPKXKKVYTSYENDL 1078
QY 612 VKE-----RKILSGY-----IVEIEDETEGLKEVINDRYDMLNI 665
Db 1079 LKKVTDGTGVTSDYDSEGLRVKQVYNSANSTEAKPVFTEYQYSGHRLKAINAKKETVY 1138
QY 630 NIDKDI-----RKILSGY-----IVEIEDETEGLKEVINDRYDMLNI 665
Db 1139 SYDADKKTLLMTQPNRKYQVGYNEAGNPQIOWIDDAEGLKITTNTKYEYNNV 1190

RESULT 15
CAGA_HELPJ
ID CAGA_HELPJ STANDARD; PRT; 1167 AA.
AC Q3ZLTI;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytotoxicity associated immunodominant antigen (120 kDa protein)
DE (CAGA pathogenicity island protein 26).
GN CAGA OR CAI OR CAG26 OR JHP0495.
OS Helicobacter pylori 399 (Campylobacter pylori 399).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
```

```
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir B.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Gullid B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Metzberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:175-180(1999).
CC -!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
CC OR FUNCTION OF THE CYTOTOXIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.iab-sib.ch/announce/
CC or send an email to license@iab-sib.ch).
CC
CC EMBL; AE001483; AAD06073.1; -.
DR PIR; B71924; B71924.
DR InterPro; IPR005169; Caga.
DR InterPro; IPR004355; IVSec_caga.
DR Pfam; PF03507; Caga; 1.
DR PRINTS; PR01553; TYPE4SSCAGA.
KW Antigen; Complete proteome.
FT DOMAIN 246 249 POLY-THR.
FT DOMAIN 882 889 POLY-ASN.
SQ SEQUENCE 1167 AA; 129729 MW; FD5E6B81CEBD0F2 CRC64;

Query Match 4.2%; Score 158.5; DB 1; Length 1167;
Best Local Similarity 20.5%; Pred. No. 1.4;
Matches 167; Conservative 138; Mismatches 297; Indels 213; Gaps 44;

QY 3 KOENPLLNESSSSQCGLL---GYFSDL-NP-----QAPWVTSSTTGDLSPSSELE- 51
Db 480 KADKALDREAKTTLGGSUKHGVNMFVYSNPKYTNASKSPDKGVGATNG-----VSHLEA 534
QY 52 -----NIPSENYFQSAIWGFIKVKSDYTPATSDNHVTWVDD-----QEVIN 98
Db 535 GFSKVAVENLPLNLAITSV---VRQDLEDKLIAGLSPQEAANKLVDFLSSNRELVG 590
QY 99 KASNNKTRLE---KGRLYQIKIQYQRENTEKGLDFKLYWTDSONKKEVISSDNLQLE 155
Db 591 KALNFNKVAEAKTNGYDEVK---QAQKOLEKSL-----KKRERLEKDVAK--N 635
QY 156 LKQKSSNRKRKSTISAGTVPDR-----DNDGIPDSLEVEGYTVDVKNKRTFLSPWISNI 210
Db 636 LESKSGNKNKMEAKSOANSQKDEIFALINKEANDARAI-AYAQLNKGIKRELSDKLENI 694
QY 211 HEKKGGLTKYKSPKSWSTASDPYSDPEK-----VTGRI-DKNVSPEARHPLVAAYPIV 262
Db 695 N--KOLKDFSKSFDEFKNGKN--KDFSKASETLKALGSKVDLGINPEWIS----- 741
QY 263 HYVDMENIILSKNEQSTQNTDSETRITISKNTSTSRTHSEVHGNAEVAHSPFDIGGSVSA 322
Db 742 --KVENLNAALNEFKNGKDKFSKVTAQK--SDLENSIKDVIINQKITDKVDNLNQAVSV 797
QY 323 GFSNSNSTVALDHSLSLAGERTWAETMGLNTADTARLNANIRVYNTG-TAPIYVNLPTT 381
Db 798 AKATGDFSGV-----EQALADLKNFSKEQLAQQAQKNEHDFNTGKNSALYQ----- 842
QY 382 SLVGLKQNTLATIKADENQSLQILAPNNYPSKNLAPI--ALANQ-----KDS 428
Db 843 SVKNGVNGTLV-----GNGLSKAEATT---LSKNFSDIKKELNAKLGNFNNNNNGLNS 894
QY 429 STPIWNTNQFLEKTKQLBLDQDQVGNIAATNFENGVRVDTGSGNWEVLPQIQTET 488
```


Db 895 TEPI---YTQVAKVKAKIDRL--DQIASGLG-----DVQOASFLLRHDKV- 937
Qy 489 ARIIFNGKDLNLVERRIAAVNPSPLETTKPD-----TLXEALKIARFNEPENGNLQYQG 544
Db 938 -----DDLs---KVGLSANHEPIYATIDDLGGPPLXRHDKV----- 971
Qy 545 KDI TEFDNFQOTSQNIKN---QJAEINAT---NIYTVLDKIKLNAKMMILIRDKRPHY 598
Db 972 DDLSKVGLSREQKLTQKIDNLNQAVSEAKASHFDNLDQMIDKLDSTKKVW----- 1023
Qy 599 DRNNIAGADESVVKEAHREVINGSTEG---LLLNl-DKDIRKILSGYIIVEIEDTEGLKE 654
Db 1024 ---NLVYESAKVPTLSAKLDNATNSHTINSNVKNGTINEKATCMLTQ-KNSEWLK- 1078
Qy 655 VINDRYDMLNlSSLRQDGKTFIDF-----KKYND--KLPLYISNP-----NYKVV 698
Db 1079 LVNDKIVAHNVGSAPLSAYDKIGFNQKNMKDYSDFKSTRLSNAVKDIKSGFVQFLTNI 1138
Qy 699 YAVTKENTlINPSENGDTSTNGIKILlFSKKGYE 733
Db 1139 FSMGSYSLMKASVHGKNTN-----TKGGFQ 1165

Search completed: May 3, 2004, 19:37:04
Job time : 13.2539 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 34.6261 Seconds
(without alignments)
6697.424 Million cell updates/sec

Title: US-09-848-909A-13
Perfect score: 3772
Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKKILFSSKGYEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25.*

```

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rviro.*
16: sp_bacterioph.*
17: sp_archaeap.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	3681	97.6	741	2	Q937W2	Q937W2 bacillus an
2	3681	97.6	743	2	Q937W3	Q937W3 bacillus an
3	874.5	23.2	876	2	Q32739	Q32739 clostridium
4	874.5	23.2	879	2	O06498	O06498 clostridium
5	872.5	23.1	876	2	Q9KH41	Q9KH41 clostridium
6	868	23.0	875	2	Q46221	Q46221 clostridium
7	810	21.5	721	2	O86171	O86171 clostridium
8	720	19.1	775	2	O844J8	O844J8 bacillus th
9	296.5	7.9	225	2	Q8XYK2	Q8XYK2 bacillus an
10	222.5	5.9	1881	16	Q8RGK2	Q8RGK2 fusobacteri
11	215.5	5.7	3468	5	Q8II04	Q8II04 plasmodium
12	202	5.4	4688	16	Q9PQ08	Q9PQ08 ureaplasma
13	200.5	5.3	2849	5	Q8IHY4	Q8IHY4 plasmodium
14	198.5	5.3	1904	5	Q8IKS2	Q8IKS2 plasmodium
15	197.5	5.2	2940	5	Q8IHP9	Q8IHP9 plasmodium
16	196	5.2	1404	2	Q8RJN9	Q8RJN9 mycoplasma

17	193.5	5.1	675	16	Q897H6	Q897H6 clostridium
18	192.5	5.1	2771	5	Q26216	Q26216 plasmodium
19	191.5	5.1	2444	16	Q81GX1	Q81GX1 bacillus ce
20	190	5.0	9904	16	Q8NW06	Q8NW06 staphylococ
21	189.5	5.0	2529	16	O25579	O25579 helicobacte
22	189	5.0	3248	5	Q8ICP9	Q8ICP9 plasmodium
23	188	5.0	2792	5	Q8I4R2	Q8I4R2 plasmodium
24	187.5	5.0	1301	5	Q8WSK5	Q8WSK5 plasmodium
25	187.5	5.0	1681	5	Q8I548	Q8I548 plasmodium
26	186.5	4.9	1125	16	Q38PM9	Q38PM9 mycoplasma
27	186.5	4.9	2235	5	Q8IB09	Q8IB09 plasmodium
28	186	4.9	1859	5	Q8IC27	Q8IC27 plasmodium
29	186	4.9	3504	5	Q8IL45	Q8IL45 plasmodium
30	186	4.9	4361	5	Q8IFP4	Q8IFP4 plasmodium
31	183	4.9	3203	5	Q7YWE7	Q7YWE7 plasmodium
32	183	4.9	3203	5	Q7YWE6	Q7YWE6 plasmodium
33	183	4.9	6713	16	Q39U54	Q39U54 staphylococ
34	183	4.9	6713	16	Q331R6	Q331R6 staphylococ
35	183	4.9	10061	5	Q8I3Z1	Q8I3Z1 plasmodium
36	182.5	4.8	1113	5	Q8I5Q0	Q8I5Q0 plasmodium
37	182.5	4.8	1301	5	Q8IHQ2	Q8IHQ2 plasmodium
38	182.5	4.8	2049	5	Q8IDA9	Q8IDA9 plasmodium
39	182	4.8	2747	5	Q9BJX9	Q9BJX9 plasmodium
40	181.5	4.8	1263	5	Q8IK49	Q8IK49 plasmodium
41	181.5	4.8	1387	5	Q9GZ76	Q9GZ76 plasmodium
42	181.5	4.8	3080	5	Q7YWF0	Q7YWF0 plasmodium
43	181	4.8	1151	5	Q8ILU0	Q8ILU0 plasmodium
44	181	4.8	1302	2	Q49547	Q49547 mycoplasma
45	181	4.8	2269	5	Q26223	Q26223 plasmodium

ALIGNMENTS

RESULT 1

Q937W2 PRELIMINARY; PRT; 741 AA.
AC Q937W2
DT 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
DE Pag protein (Fragment).
GN PAG.
OS Bacillus anthracis.
OG Plasmid PX01.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ferrara;
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
RA Pasarella A., Francia M., Ciuchini F.,
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus anthracis vaccine strain 'Carbovap'";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ413937; CAC93935.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax_toxinB.
DR Pfam; PF03495; Anthrax_toxB; I.
DR PRINTS; PRO1391; BINARYTOXINE.
KW Plasmid.
FT NON_TER 1
FT NON_TER 741
SQ SEQUENCE 741 AA; 83153 MW; C7F95820B73065C0 CRC64;

Query Match 97.6%; Score 3681; DB 2; Length 741;
Best Local Similarity 99.3%; Pred. No. 5.3e-167;
Matches 717; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENLLNESSSQGLLGYFFDLNFOAPWVWTSSTTGDLSIPSELENIPSENOYF 60

```

Db      20 EVKQNRLLNESSESSQGLLYFSDLNFAQPMVVTSTTGDLSIPSSLENIIPSENOYF 79
QY      61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db      80 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 139
QY     121 QRENTEKGLDFKLYWTDSONKKEVITSSDNLOLPELKQKSSNSRKRSTSGPTVPDRDN 180
Db     140 QRENTEKGLDFKLYWTDSONKKEVITSSDNLOLPELKQKSSNSRKRSTSGPTVPDRDN 199
QY     181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTASDPYDFEYVT 240
Db     200 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTASDPYDFEYVT 259
QY     241 GRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db     260 GRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 319
QY     301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERETWAEITMGLNTADTARL 360
Db     320 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERETWAEITMGLNTADTARL 379
QY     361 NANIRVYNTGTAPIYNNVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYTPSKNLAPIA 420
Db     380 NANIRVYNTGTAPIYNNVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYTPSKNLAPIA 439
QY     421 LNAQKDSASTPIIMVYNNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480
Db     440 LNAQKDSASTPIIMVYNNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 499
QY     481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 540
Db     500 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 559
QY     541 QYQKDIETEDFDFNDQSTQNIKNQLAELNATNITVLDKIKNAKNNILIRDKRHFYDR 600
Db     560 QYQKDIETEDFDFNDQSTQNIKNQLAELNATNITVLDKIKNAKNNILIRDKRHFYDR 619
QY     601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYVEIETEDGLKEVINDRY 660
Db     620 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYVEIETEDGLKEVINDRY 679
QY     661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYNNVAVTKENTIINPSENGDTSTNG 720
Db     680 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYNNVAVTKENTIINPSENGDTSTNG 739
QY     721 IK 722
Db     740 IK 741

```

RESULT 2

```

Q937W3 ID Q937W3 PRELIMINARY; PRT; 743 AA.
AC Q937W3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pag protein (Fragment).
GN PAG.
OS Bacillus anthracis.
OC Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1].
RP SEQUENCE FROM N.A.
RC STRAIN=Carbosap;
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
RA Fasanella A., Francia M., Ciuchini F.;
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus Anthracis vaccine strain 'Carbosap'.";
RL Submitted (SEF-2001) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AJ413936; CAC93934.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxB; I.
DR PRINTS; PR01391; BINARYTOXINB.
KW Plasmid.
FT NON_TER 1 743
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5D5A6 CRC64;

Query Match          97.6%; Score 3681; DB 2; Length 743;
Best Local Similarity 99.3%; Pred.No.5.3e-167;
Matches 717; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQNRLLNESSESSQGLLYFSDLNFAQPMVVTSTTGDLSIPSSLENIIPSENOYF 60
Db 22 EVKQNRLLNESSESSQGLLYFSDLNFAQPMVVTSTTGDLSIPSSLENIIPSENOYF 81
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 82 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 141
QY 121 QRENTEKGLDFKLYWTDSONKKEVITSSDNLOLPELKQKSSNSRKRSTSGPTVPDRDN 190
Db 142 QRENTEKGLDFKLYWTDSONKKEVITSSDNLOLPELKQKSSNSRKRSTSGPTVPDRDN 201
QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTASDPYDFEYVT 240
Db 202 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTASDPYDFEYVT 261
QY 241 GRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 262 GRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 321
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERETWAEITMGLNTADTARL 360
Db 322 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERETWAEITMGLNTADTARL 381
QY 361 NANIRVYNTGTAPIYNNVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYTPSKNLAPIA 420
Db 382 NANIRVYNTGTAPIYNNVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYTPSKNLAPIA 441
QY 421 LNAQKDSASTPIIMVYNNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480
Db 442 LNAQKDSASTPIIMVYNNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 501
QY 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 540
Db 502 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 561
QY 541 QYQKDIETEDFDFNDQSTQNIKNQLAELNATNITVLDKIKNAKNNILIRDKRHFYDR 600
Db 562 QYQKDIETEDFDFNDQSTQNIKNQLAELNATNITVLDKIKNAKNNILIRDKRHFYDR 621
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYVEIETEDGLKEVINDRY 660
Db 622 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYVEIETEDGLKEVINDRY 681
QY 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYNNVAVTKENTIINPSENGDTSTNG 720
Db 682 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYNNVAVTKENTIINPSENGDTSTNG 741
QY 721 IK 722
Db 742 IK 743

RESULT 3
O32739
ID O32739 PRELIMINARY; PRT; 876 AA.

```



```

RN  SEQUENCE FROM N.A.
RP  STRAIN=NCIB 10748;
RX  MEDLINE=94041637; PubMed=8225592;
RA  Perelle S., Gibert M., Boquet P., Popoff M.R.;
RT  "Characterization of Clostridium perfringens iota toxin genes and
RL  expression in Escherichia coli.";
RN  Infect. Immun. 61:5147-5156(1993).
RN  (2)
RP  SEQUENCE FROM N.A.
RA  STRAIN=NCIB 10748;
RC  Popoff M.R.;
RL  Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; X75562; CAA51960.1; --
DR  F01; 140862; I40862.
DR  HSSP; P13423; IACC.
DR  GO; GO:0005576; C:extracellular; IEA.
DR  GO; GO:0015070; F:toxin activity; IEA.
DR  GO; GO:0009405; P:pathogenesis; IEA.
DR  InterPro; IPR003896; Anthrax toxins.
DR  Pfam; PF03495; Binary toxin; I.
DR  PRINTS; PR01391; BINARYTOXINB.
KW  Signal.
FT  SIGNAL.
FT  CHAIN.
SQ  SEQUENCE 34 38 POTENTIAL.
      212 875 IOTA TOXIN COMPONENT IB.
      SEQUENCE 875 AA; 98468 MW; C9AE092CD3818921 CRC64;

Query Match      23.0%; Score 868; DB 2; Length 875;
Best Local Similarity 31.1%; Pred. No. 2,7e-33;
Matches 222; Conservative 130; Mismatches 279; Indels 150; Gaps 31;

QY  1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSELENIPSE-NQY 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      35 DTNQKEBITNENTLSSNGLMGYFADHFHFKOLELAPIKNGDLKPEEKVLDLTDNNS 94

QY  60 FQSAIWGGITKVKKSDEYFATSAADNHTVMVDQDEVINKASNNKIRLEKGLYQIKITQ 119
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      95 IKSIWTGRIIPSEGEYILSTDR-NDVLNQINAGDIK----TLKVNKMGQAYNIRIE 150

QY  120 YQREN-----PTEKGLDFKLYWTDQSNKKEVISSDNLOLPELKQKSSNRKRSTS 170
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      151 IQDKNLGSDNLVSP-----KLYW-ELNGNKTVIPENLFFRDYKIDEND----- 195

QY  171 AGTVP-----DRNDGIPDSLEVEGYTVVKNKRTFLSPMISNIH 211
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      196 --PFIPNNFFDVRFFSAWEDEDLTDNDNIPDAYEKNGYTI-----KDSIAVKNDSFA 249

QY  212 EKKGKTKYKSSPEKWTASDPYSDFEKVTGRIDKNVSPPEARPLVAAPYIVHVDMENIL 271
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      250 E-QGYKKYSSYLENTAGDPYDYQKASGSDKALKLEARDPLVAAPYVVGNGHENLII 308

QY  272 SKNEDQSTQNTDSEETISKNTSTSTSTSEVHGNNAEVHASFDDIGGSVSAFSSNSST 331
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      309 STNEHASS-----DQKTVSRATTSKTDANTV-----GVSTISAGYQNGFTGN 351

QY  332 VAIDHS-----LSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVNLPTSL 383
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      352 ITTSYHTTNDSTAVQDSNGESNWTGSLNKESAYINANVYNTGTAPYKVTPTNL 411

QY  384 VLKGNQTLATIKADENQOLGQILAPNNYFSKNLAPIALNAQKADASTPTIMYNGFLE 443
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      412 VL-DGETLATIKADNQIGNLSPNETYPKGLSPLALNTMQFNARLIPINDYQKKLD 470

QY  444 KTKQLRLDQVYGNATYVNFENGRVVDTSNWSSEVLPOIQTETARIIFNGKDLNLYER 503
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      471 SGQIKLETTQVSGNGYGTGN-SGQI-ITEGNSWNSYISQIDSVSASIIID-TGSQTER 527

QY  504 RIAAVNPSPLETTKPDMTLKEALKIAGFNPNPENGLOQ-OGKDLITE--FDNFQDQTSQ 560
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      528 RVAAKEGPNEDKT-FEITIGEAIKKAFASTK-NGELLYFNGPIDESCVELIFDNTSE 585

QY  561 NIKNQIAELNATNIYVLDKILNAQONILIRDKRF--HYDR-NNIYVAVGADSVVKEAHR 617
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 586 IIKEQLKYLDDKKIYV----KLERGNLILKVPSYFTNPFDEYNFP--ASWSNIDTKNQ 639
QY 618 EVINSSTEG------LLNIDKIDKIRKILSGY-----IVEIEDTEGLKEVIND 658
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 640 DGLQSVANKLSGETKIIPMSKLPKRYVFGYSGKDPSTNSITVNIKSKBOQTDYLPV 699
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 659 RYDMLNIS-----SLRQDGKTFIDFKYND--KLPLVISPNPKV----- 696
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 700 EKDYTKFSYEFETTKGDSSDIEITLSSGVIFLDNLSITELNSTPILKEPEIKVPSDQE 759
QY 697 -----NVYAVTKENTIIINSENGDTSTNGI 721
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 760 ILDAHNKYADIKLDT-----NTGNTYIDGI 785

RESULT 7
O86171 AC O86171 PRELIMINARY; PRT; 721 AA.
ID AC O86171;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE C2 toxin (Component-II).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=96323874; PubMed=9659689;
RA Kimura K., Kubota T., Ohishi I., Isogai H., Isogai B., Fujii N.;
RT "The gene for component-II of botulinum C2 toxin.";
RL Vet. Microbiol. 62:27-34(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=96184657; PubMed=8645309;
RA Fujii N., Kubota T., Shirakawa S., Kimura K., Ohishi I., Moriishi K.,
RA Isogai E., Isogai H.;
RT "Characterization of component-I gene of botulinum C2 toxin and PCR
RT detection of its gene in clostridial species.";
RL Biochem. Biophys. Res. Commun. 220:353-359(1996).
DR EMBL; D88982; BAA32537.1; --
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxin; I.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 721 AA; 80515 MW; 44C8153AC749D5F2 CRC64;

Query Match      21.5%; Score 810; DB 2; Length 721;
Best Local Similarity 35.3%; Pred. No. 1.2e-30;
Matches 223; Conservative 117; Mismatches 228; Indels 64; Gaps 25;

QY 3 KOENRLNES--SSQGLGYFSDLNFPQAPMVVTSSTGDLSPSELENIPSENQYF 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      5 KFENSVKSNKRYNTINGLMGYFEN-DFNLIISITPLDGNLTFKEDINSILG-NKII 62

QY 61 QSAIWSGFIKVKKSDEYFATSAADN-HVTVMVDQDEVINKASN-SNKIRLEKGLYQIKI 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      63 KSAEWIGLIKPSITGEYILSTNSPNCVEL--NGEIFNLSTNTVNLIQGVYDIRI 119

QY 119 -QYORENPTKGLD-FKLYWTDQSNKKEVISSDNLOLPELKQKSSNRKRSTSAGTV- 175
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      120 EQLMSNQLLNKYGKILYWTSDIIEIIPSEVL-----LKNYSNTNKSFTIPNNTLF 175

QY 176 -----PDRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEK 225
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      176 SNAKLKANARDDTRDGIPEWEINGYTVVKNQKAVAWDDKFAAN-----GYKYVSNPFK 230

```

```
QY 226 WSTADSPYDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEPOSTONTDSE 285
D 231 PCTANDPYDPEKVSQIDPSVSWARPMISAYPIGVQWERLVVSKSE-----TITGDS 286
QY 286 TRTSKNTSTSTHTSEVHGNVAFVAFDGG-----SVSAGFNSNSNSTVAIDHSL 338
D 287 TKSMKSTSHSTNTNTV--GAEVSGSLQAGGIPPFVFSMSASANSYHTWQNTSTVDDT- 343
QY 339 SLAGERTWAEWTGLTADTARLNANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADE 398
D 344 --TGE-SFSGSLNIGSAYINPIRYNTGTAPVYNTPTTIVIDK-OSVATIKGE 399
QY 399 NOLSOILAPNYYPSKNLAPIALNAOKDASSTPIITMNYNQFLELEKTKQLRLDQVGN 458
D 400 SLIGDYLNPGGTYPIIGEPMALNTMDOFSSSLIPINYNQLKSIDNGGTVMLSQFTGN 459
QY 459 IATYFENGVRVVDGNSWSEVLPOIQTETARII--FNGKDLNVERIAAVNPSDPLET 516
D 460 FAKYN-SGNLIVD-GNNWGPVLGTTKTASLTLSFSGQTQVA--VVAFNSDPEBK 514
QY 517 TKPDMTLKEALKIAFGFNEPNNGNLOYQOGKDI--EFDNFDOOTSONIKNQLAELNATN 573
D 515 T-PKLTLEQALVAFKALEKNGKVFHGLEISKNEKIQVFLDSNTNDFENQLKNTADKD 573
QY 574 IYVLDKIKLNAMNILLIRKPHYDRNNIAV 605
D 574 IMHCI--IKRN--MNLVKVITFKENISSINI 601

RESULT 8
Q844J8
ID Q844J8 PRELIMINARY; PRT; 775 AA.
AC Q844J8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Viplac.
GN Viplac.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y., Chen J., Yang Y.
RT "Cloning of vip1A(C) and vip2A(C) from Bacillus thuringiensis HD201.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY245547; AA08614.1;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxins.
DR Pfam; PF03495; Anthrax toxin.
DR PRINTS; PR01391; BINARYTOXIN.
DR SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;

Query Match 19.1%; Score 720; DB 2; Length 775;
Best Local Similarity 29.0%; Pred. No. 2.5e-26;
Matches 223; Conservative 133; Mismatches 294; Indels 120; Gaps 30;

QY 4 QENRLNESSSQGLGYFSDLNFPQPMVVTSTTGDLSPSSLEN--IPSENQVFO 61
D 42 QKQK---QKMDRKGLLGYFKGDF-SNLTWFAPTEDNTLYDQGTANKLDDKQKQYQ 97
QY 62 SATWSGFIKVKSDYVTFATPSADNHYTMVDDQEVINKASNNKRLKGRLYQIKQYQ 121
D 98 SIRWIGLIQSKGKDFNLSEDEQAIIEIDGKIIISNKGEKQVHLEKELVPIKIEYQ 157
QY 122 RENPTEKGLD-----PKLWYTSQNKKEVISSDNLQLPDLKQKSN----- 162
D 158 SD--TKFNIDSKTFKPKLFRIDSONQSQKDELNPFENFKESREFLAKASKTNPFM 215
QY 163 SRKKRSTAGPTVPRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKS 222
```

```
216 QRMKRDDED---TDTGDSIPDLWEANGYT--IQNK--VAVKW-DDKFAQGYKYKLS 267
223 BEKWSFASDPYDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEPOSTONT 282
268 PIQAHVGPYDWEKAAGDIPKSWAATRNPLVAFPSINVDMERKMLSKDNLNLS---- 323
283 DSETRTISKNTSTSTHTSEVHGNVAFVAFDGG-----SVSAGFNSNSNSTVAIDHSL 339
324 -----NSAEHNSNSTYVANSB-----GASIEAGFGPKGFSFGVSANYQHTET 366
340 LAGERTW---AETMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIK 395
367 VGSD--WGNKSNTEOFNSASAGYLNANVYNNVGTGGIYDAQPTTSFIL-QDSIATIT 423
396 ADENQISQILAPNYYPSKNLAPIALNAOKDASSTPIITMNYNQFLELEKTKQLRLDQV 455
424 AKSNATALSTPSGDRYPASK-EGISLKTWDDFNPHITLTKPQLDAVLNNEVIKINTDQT 482
456 YGNIAFYFENGVRVVDGNSWSEVLPOIQTETARII--FNGKDLNVERIAAVNPSDPLE 515
483 DGRYGIIGVDG---KAEIGDRWSPIIDEIKGRTASIIIDPADGKALETRIAAKDYKNPED 539
516 TTKPDMTLKEALKIAFGFNEPNNGNLOYQOGKDI--ITEFDNF-----FDOOTSON 561
540 KT-PSLTIKEGKLIAYPESISEDKGILFVEYKNDGKVTKKQLSEENIMPLYLDETSKE 598
562 IKNQLAELNATNYVLDKIKLNAMNILLIRKPHYDRNNIAVGADESIVVKEAHEVIN 621
599 FERQLSDGSAGLY---DKLTPKKNITIR-----LATVLGDDDFDSAYFWENATW 647
622 SSTEGLL-----LNIDKDIRK-----ILSGYIVE--IEDTEGLKEYI---NDR 659
648 SDKFGNLRGLSLAIPOESKYTIIPKDKVKPNYDYLITGYIKHDFTTDNESIGIVAFTKON 707
660 YDMLNITSSLRQDGKTIFIDKKNYDKLPLYSNPYKVNKYVANTKENTIN 709
708 FEMNMGTSIFSQNSGGFEKFTIKTQ-NISG-DYILDSIQLMKRNDVN 755

RESULT 9
Q8KYK2
ID Q8KYK2 PRELIMINARY; PRT; 225 AA.
AC Q8KYK2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protective antigen-related protein, (pXO1-111).
GN BXA0163.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STEAIN-A2012;
RX MEDLINE=22061436; PubMed=12004073;
RA Read T.D., Salzberg S.H., Pop M., Shumway M., Umayam L., Jiang L.,
RA Holtzapflee E., Busch J.D., Smith K.L., Schupp J.M., Solomon D.,
RA Keim P., Fraser C.M.;
RT "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in
RT Bacillus anthracis.";
RL Science 296:2028-2033 (2002).
DR EMBL; AE011130; AAM26108.1;
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 225 AA; 25402 MW; 2E121BES4295F9C8 CRC64;

Query Match 7.9%; Score 296.5; DB 2; Length 225;
Best Local Similarity 37.2%; Pred. No. 6.4e-07;
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;

QY 566 LAELNATNYVLDKIKLNAMNILLIRKPHYDRNNIAVGADESIVVKEAHEVINSTE 625
```



```

Db 1 MESLGINNIYNALDRKLNKNNILVRDP-YHYDNNGNIVGDDSYLKNAYKQILNWSDD 59
QY 626 GLLNIDKIDKILSGYIVEIEDTE-----GLKEVINDRYDMLNLSLRODQ 672
Db 60 GVSMLDEVDQALSGYMLQKKPSNHLTNSPVITLAGKDSGVGLYRVLS-----DG 113
QY 673 KTFIDFKYNDKPLIYISPNKYKVYAVTKEN-TIINPSNGDTSTNGIKKILIFSKKG 731
Db 114 TGFLDFNKFDENRSLV-DPGDDVYVYAVTKEDFNNAVTRDENGIA-NKLKNTLVLSGKI 171
QY 732 YEI 734
Db 172 KEI 174

RESULT 10
Q8RGK2 PRELIMINARY; PRT; 1881 AA.
AC Q8RGK2;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE Hemolysin.
DE Hemolysin.
GN FN0291.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RC MEDLINE=2186394; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen M., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyripides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AEO10541; AAL94497.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008638; Haemagg act.
DR InterPro; IPR000508; Peptidase_S26.
DR InterPro; IPR001452; SH3.
DR Pfam; PF05860; Haemagg_act; 1.
DR Pfam; PF00018; SH3; 1.
DR PROSITE; PS00761; SPASE_I_3; 1.
KW Complete proteome.
SQ SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;

Query Match 5.9%; Score 222.5; DB 16; Length 1881;
Best Local Similarity 20.5%; Pred. No. 0.034;
Matches 173; Conservative 133; Mismatches 303; Indels 233; Gaps 36;

QY 30 QAPMVVTS-----STTG-----DLSIPSGELE--NIPSENV-FOSAIWSGFI 69
Db 246 QAPWATKGDVVISKGVYLDQTQAKRDIKISSTEIGSKLLAENAINIKSGKTSNG 305
QY 70 KVKSDYEYFATSADNHVTWVDDQEVN-----KASNSKIBLEKGR 112
Db 306 QIRANNITINGNVDSNLIFTNKDITISGLKNSGVSSSNLVNKEIENSNKVYVEE-K 364
QY 113 LYGIKI-----QYQRENPEKGLDFKLYWT--DSQNKKEVISSDNLQLPELKQKSS 161
Db 365 LSSTKITNLGNLSAKIEKTNFNSGKLFSGKITAKDFKNGEV-SSENLITTNLENSK 423
QY 162 -----NSRKGSSTAGTPVPRDNDG----- 182
Db 424 INVKENINSIVNKTNAETTSKNTLNLDNRGNTIINNVSsgVIANGKLLVGNITN 483

```

```

QY 183 ---IPDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKV 239
Db 484 SQNLNTATVQCKTLDDIKNINSGSKILSDNLTLDIPSSGNISAKVITTOELINSGEII 543
QY 240 TGRIDKNVSPERHPLVAAPYPIVHYDMEN-----ILLSKQEDQSTQNTDSTRISKNT 293
Db 544 SNLNNNNNNKSNIFVGNLKI SNLNNNSGVIEGLELNTSIENGTITKNKLTSQL 603
QY 294 STSRTHTSVHGNAEVHASFDDIGGSVAGFSNSSTVA---IDHSLSLAGERATWAEFM 350
Db 604 NNKK-----NTANVNAAGFLDVHNKISS-VGNIKAITMKTNNLDSNGNIL-----IN 648
QY 351 GLNTAD-----TARLNANIRYVNTGTAPITVNVLPFTSLVLGKQOTLATIKADENQLSQ 403
Db 649 SLTTAENINKSGITAKNISQNLVNSGS-----VISDNITVA-----DN 687
QY 404 ILAPNNYPSKNLAPIALNAQKQDASPTITMNYNOFLELEKTKQLRLDQDVYGNIAATYN 463
Db 688 ITWNNIPANEK-----ISADKISNSNK-----LVAKNTEITKLTND---GNIVVK- 730
QY 464 FENGRRVDTGNSWSEV-----LPQIQTARIIFNGKOLNIVERRIAAVNPSDPLE 515
Db 731 -ENLKAKDITNSNTIKVGENLNTDKLQSKTLIAKNINIEKSLNININGKITSLNAINITS 789
QY 516 TTKPDMTLKALKIAPGFNEPENGNYQYQKGDITEPDFENFDQQTSONIKQLAELNATNY 575
Db 790 DIKNNNGITQAIK-----NINIKTSNDLKDGKYTANDSLNINAKSLNNGN--- 836
QY 576 TVLD---KIKLNAMNILLIRKFRHYDRNNAIV-----GADESVVYKEAHREVINS--T 624
Db 837 --LENDGKIKFNLGTNLNNKISSSNLNTANETISNNGVNSIIGSEANLTITANSLNK 894
QY 625 EGGLE-----NI-DKDIRKILSGYIVEIETEGLEKVINDRYDMLNLSLRSQ 670
Db 895 EGNLLFGEGIENKLTGNTITGVTSISLGKUKIEADVNDKHIISDNDLIDVNSITN 954
QY 671 DGTFTF-----IDFKK--YNDKLLPLYSNPYKNVYAVYKENTINPSENGDTSTNGIK 722
Db 955 KGLLYSTNNMKVDFKENFLNDKAEIYSSG-----DITIN-SENG-TFTNRVG 999
QY 723 KI 724
Db 1000 DI 1001

RESULT 11
Q8II04 PRELIMINARY; PRT; 3468 AA.
AC Q8II04;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pfl_0371.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22285705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung B., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Arguioli S.,
RA Pette M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C., R.W.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).

```

DR EMBL; AE014841; AAN35955.1; -;
KW Hypothetical protein.
SQ SEQUENCE 3468 AA; 411913 MW; 506F7D62999BA7B1 CRC64;

Query Match 5.7%; Score 215.5; DB 5; Length 3468;
Best Local Similarity 19.4%; Pred. No. 0.16; Indels 275; Gaps 40;
Matches 162; Conservative 134; Mismatches 266;

QY 55 SENQYFQSAIGWGFIVKKSDEYTFATSDNHTVMWVDDQEV---INKASNSKIRLEKG 111
DB 160 SENEV-----TKNDKYNSTNSGY-----NKEIEFLKNAIEHSLKTEKN 202

QY 112 RLYQIK-----IOYRENPTEKGLDFK---LYW-----TDSN-----KKEVISDNL 151
DB 203 INLLKCDSDNNDLFEKKNVHLNNDYSEDLFHDNRVNNQNNLKDIIKDYIHDNI 262

QY 152 QLPKQKSSNSRKRKRSAGTVPDRDNDGIPDSLEVEGYTVDV-KNKRTFLSPWISNI 210
DB 263 MI-----ESNDNRNDKNCFKNTNPNYERHIIIVDKTKNGNNNNNSNNFDS--ISNI 315

QY 211 HEKGLTKYKSPKWKSTASDP-----YSDPE-KVTGRIDKNVSPKARHPLVAAYPIVHV- 264
DB 316 NKK-----ISYPNNVNSSEDKTLNQINLDMSLSSDSLXNAYSFYSL 360

QY 265 ---DMENILSKNEPOSTQNTD--SETRTISKNTSTRTSTSEVHGNAEVAH---SFFDIG 317
DB 361 EKXNNIGMGNHKNWMEYRVIDEDINISKFSNISLHDSEKNDNNEFNSLCSP--- 416

QY 318 GVSAGFNSNSSTVAIDHSLSLAGERTWAEFMGL-----NTADTARLANIRY----- 366
DB 417 -----NSSKCKVDEMITQVGNRKVYKVTMDVANDNYMKNSTECISNMESNHFKNQ 468

QY 367 -VNTGTAPIYVLPVTSVLGKQ-----TLATIKADE--NOLSOILAPNN--- 409
DB 469 KINSKEDANNLNSHSHILANNKIGQVNSLSDYLSIKQESNMNLSNNEALNINV 528

QY 410 -----YYPKSLAPIALNAQKASSTPITWYNO---FLELE 443
DB 529 TNNNGSSNNKNSNVYKTSQVYVNPNDIN--NMNHNLSYMKNSNNLNTSNGFKKIP 586

QY 444 KTKQL--RLDTQ-----VYGN-----ATYVNFGRVVDVTGNSWSEVLQ 483
DB 587 KKNKII-SNIDFDNNIFKSVKENVKNQESNINHOFVEKYNINDEINIKN-----NNIENN 642

QY 484 IQETAR-----IIFGKDLNLVERRIAANVPSDPLETTK----- 518
DB 643 TQNTTCNNFINTDDIINKKIKIYKIDSISLLDLSLKLSSINSSIMDRYTKNVE 702

QY 519 -----PDMTLKEAL-----KIAFGNEPGLQYQKDIETEFDFNQ 557
DB 703 EKFLDDVILDDSI-FATSNELLOHSNYTTTTHIFDNNNNNNNNINQEKDLFQNDYN--K 759

QY 558 TSONIKNQLAELNATNIYTVLDIKLNAKNILIRDRFHYDRNINAVGADSESVVKEAR 617
DB 760 ETYNNIMLSENNAELF-----KISYCNOLVLGKNEIILD--R 798

QY 618 EVINSSTEGLLANIDKDIRKILSGYVEIEDTEGLKE-----VINDRYMLNLS-- 667
DB 799 NVENSKE-----QVEEYQNKEDIKLYHKDDNIIIDNNYDVKKKDC 842

QY 668 -LRQDGKTFDFKYNQ---KLPL-----YISNPYKVN-VYAVTKNTINPSEN 713
DB 843 HLKMDNQDNINKKQSKQKPKINNNECNVHNKFKINEIDHFKENTLVNESKN 899

RESULT 12
Q9PQ08 PRELIMINARY; PRT; 4688 AA.
AC Q9PQ08
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Hypothetical protein UU482.

GN UU482.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=Serovar 3;
RC MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
urealyticum";
RL Nature 407:757-762 (2000).
DR EMBL; AR002145; AAF30894.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 4688 AA; 534880 MW; B53BFAFEE1997E CRC64;

Query Match 5.4%; Score 202; DB 16; Length 4688;
Best Local Similarity 21.3%; Pred. No. 1;
Matches 174; Conservative 131; Mismatches 325; Indels 186; Gaps 40;

QY 19 LLGYVESDLNFCAPMVVTSSITGDLSPSSELENIPSENQYFQSAIWSGFIKVKKSDYET 78
DB 3699 LVDVYILD-NIHNQIDETKIFKDNHV-SKEIINPGVTMISKHGNWKSPTDTTANFEFK 3756

QY 79 FATSADNHTVMWVDDQEVINKASNSKIRLEKRLYQIKIYORENPTEKGLDFKLYWTD 138
DB 3757 IETQ-----DDNDVLNNDIATVKFKDEHNNIKQKIVRIKEN-----ND 3795

QY 139 SQNKKEVISDNLQLDELKQKSN-----SRKRSISAGTVPDRD----- 179
DB 3796 WLKGGI---DNLN--PETKYKLENIELSPLKTHNLNLSVINDKENISLITGPNVLKV 3851

QY 180 ---NGCIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWKSTASDPYSD 235
DB 3852 IQQNTTINDTQQTINVTLSGVNSK-YNGRQIKVYKDNNNVYESS---LITLQKGRKD 3907

QY 236 FEKVTGRIDKNVSPKARHPLVAAYPIVHVDMENILSKN-EDOSTQNTDSE-RTISKNT 293
DB 3908 YQLLSNLNSN-----REYRFEKIEINHISNTNNFEDEKUNGVSNTFITQTKT 3957

QY 294 STSRTHTS-EVHGNAEVAHASF-----FDIGGSVSAGFS-----NSNS---STVAID 335
DB 3958 TVQWNDSSATIVTGRVGNFNFKIKSEDKILENNQVAVFAPKETIRDTNTWLYTRPLK 4017

QY 336 HSLSLAGERTWAEFMGLNT---ADTARLANIRYVYTGTAPIYVLPVTSVLGKQNTLA 392
DB 4018 DVTSDFKEGTWADLSNSVNFKEBETTVKLVKIQVNRKPKAKNNINNSNNVILDTNISI 4077

QY 393 -----TIKADENOLSOILAPNNVYPSKNLAPALN-AOKDASSTPITWY--NOFLELE 443
DB 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQTINFTLSGVKSWVGKKILSKYKSDTSESI 4137

QY 444 KTKQLRLDTPQYVGNIAATYFNGR-----VRVDTGNSNWSEVLQIQETETARIIFNGKD 497
DB 4138 HTNEVLIESNKTQYINLLNLLKRNRTYTLIDVLDLNNVSDPFGKGNLTNSPTTTSIA 4197

QY 498 LNLVERIAAVNPSDPLETT-----KPDMTLKEA-----LKIAFGNEPGLNQ 541
DB 4198 INVNLBEISNRASTNLKSTIIKLNLDNPNVLRDQOATIVYGNKQKNGFITVSGNIK 4257

QY 542 YQGGDIETEFDFNQDQTSQNIK-NQLAELNATNIYTVLDIKLNAKNVNIILIRDRFHYD- 599
DB 4258 YLTATVLDLNFN-DKVNIVNISFNKPSIAAEN-----IGIDKSNII-----YVND 4304

QY 600 -----RNIIAAGA---DESUVKEAHEVINSSTEGLLANIDKDIRKILSGYVEIEDTE 650
DB 4305 IPKLEINNDIIVNGPIINKEIVVKNANQ--NNIDVDLGLQINPKIAHLR-FIAKFKSTN 4361

QY 651 GLKEVIND--RYDMLNLSL-RODGKTFIDF-----KKY----- 681
DB 4362 -----NDIIEITVINGSLVNDGKTSIRFTLANLKANKLYSLVDVYVYLVNNSNTIVE 4415

DE Hypothetical protein UU482.

```
QY 682 NDKPLVYISPNYKVNVAVTYKENTINPSENGDTS 717
DB 4416 SNKLP-KLNNINYOIK---INKGHTII--SKNGEWS 4445

RESULT 13
Q8IH4 PRELIMINARY; PRT; 2849 AA.
AC Q8IH4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pf11_0392.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perte M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrall B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014841; AAN35975.1; -.
KW Hypothetical protein.
SQ SEQUENCE 2849 AA; 335916 MW; B5515D173D96B813 CRC64;

Query Match 5.3%; Score 200.5; DB 5; Length 2849;
Best Local Similarity 20.6%; Pred. NO. 0.64;
Matches 170; Conservative 133; Mismatches 334; Indels 189; Gaps 41;

QY 3 KQENRLNSESSESSQGLGY-YFDLNPQ--APMVVTSSTGDLSP--SSELENIPSEN 57
DB 1862 KHNYTFNKEENLKSIFKYNNNNNNNEKDDIPKSIQSSFTNEDNIAVYKNGNNMKLN 1921
QY 58 Q-YFQSAIWGFIKVKKDEYTFATSDADNHTVMVDD-----QEVINKASNNK--- 105
DB 1922 EDIQRFSHNKIKYNNKNCSDSCNIVKINYDELNDSTQTKELNEGKSNNGKAEAW 1981
QY 106 -IRLEKGRLY-QIKIQYORENTEKGLDFKLYWTDSONKKEVISEDNLQLPKQKSSNS 163
DB 1982 IIDIKNNETYPYKIDKKEKNEDEK--KNKYNNKNDNNKNTMGSSN---KSMCKNNKNS 2036
QY 164 RKRSTSGACTVPDRDNGIDPDSLEVGTVVDKVKRTFLSPWISNTHEKGLTKYKSSP 223
DB 2037 NKQKH-----IPLSVNNKY-----NKSSINKYENNIN-----KNK 2069
QY 224 EKWSTASDPYDF--EKVTGRI-DKNSGPEARHPLVAAYPIVHDM-----NILS 272
DB 2070 DKLNLVNSISKLVSQKIQELSNKISKD-----ILNFEITKIKKSKKETKNTNTN 2124
QY 273 KXEDOSTQNTDSETISQNTSTSTHISEV-HGNAEVHAFPDIGGSVSAG-FSNGSS 330
DB 2125 KNDNNNDNNNNFRINDNINSYNNLVANKVPYDINI-----DEGNVINTGPIYADGS 2179
QY 331 TVAIDHSLSLAGERWAETWGLNTA-----DTARLNA-----NIRYVNTGTAPIYVNL 378
DB 2180 TI-----YTWVNNIDTNYNNKYFDSKKNINHVPLNNVPLN-----NDL 2221
QY 379 PTTSLVLGNQQLATIKADENQLSQ-----ILAPNNYFSKULAPALNAQKDSSTPTM 434
DB 2222 LINNVILNQNNMNNLENLNTNTIGSVQPFVTCDFY--ANNIKSYLD-----PNLQ 2271

QY 435 NYNOFLELEKTKQLRLDT-----DOVYGNIAATYNFENGSRVDT-----GSN 476
DB 2272 NNNYFDNNQLLHNHNNLNNLSINYIDONYLSYNNLNCINGINIKOTCKDIVIGIPNTN 2331
QY 477 WSEVLPOIQETTARIIFNGKDLNLVERRIAAVNPSPDELE-----TTKPDMTLKALK 528
DB 2332 QNOI-PTI-ELDDTILKN--DVNLLNNNNVNVHNVVEMLNNIQTNVQKLYNDIQENLH 2387
QY 529 IAFGF--NEPNGNLOYQKGD-----ITEFPNFDQOTSONIKQLAELNATNIYTV 577
DB 2388 ITNSLHNNNNNNNNMLGGDVLLNNSYLFNINSFPNNMAYIYNNNNNNNNNNCINYL 2447
QY 578 LDKIKLNK-----WNILIRDKRFHYDRNNIAVGADESVVKEAHEVINSSTEGLLNID 632
DB 2448 LDKKEINVKNSEINWNSVLEDTNKENEGD-----MKDNKYIINK--ENNVMKND 2496
QY 633 KDIRKILSGYIVETEDTEGLKEVINDRYDMLNLSLRODGKTFDFPKYNDKLPYISNP 692
DB 2497 NNIDIISGQNTNTNANTMKERNLNRDDDIYKNNMKRNSLDFKGLDSE-----KK 2550
QY 693 NYKVVV---YAVTKENT--IINPSENGDTSTNGIKKILIFSKGYE 733
DB 2551 NKSMNIEGARKVKNTTQLLNYSEN---RKGLDMSYADKVL 2593

RESULT 14
Q8IKS2 PRELIMINARY; PRT; 1904 AA.
AC Q8IKS2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Ferlin, putative.
GN Pf14_0530.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perte M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrall B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014825; AAN37143.1; -.
DR InterPro; IPR000008; C2.
DR Pfam; PF00168; C2; 5.
DR SMART; SM00239; C2; 5.
DR PROSITE; PS00004; C2_DOMAIN_2; 2. A36470367BDEC41 CRC64;
SQ SEQUENCE 1904 AA; 224720 MW; A36470367BDEC41 CRC64;

Query Match 5.3%; Score 198.5; DB 5; Length 1904;
Best Local Similarity 19.8%; Pred. NO. 0.47;
Matches 174; Conservative 128; Mismatches 328; Indels 247; Gaps 43;

QY 3 KQENRLNSESSESSQGLGYYSFDLN-FQAPVVTSTTGDLSIPSELENIPSENQYFQ 61
DB 836 KQNN--KKQNNNNNNNNYNNSSVYQNDLLYGERVGDILL---YFELVQSKA-MK 889
QY 62 SAJWSGFIKVKKDEYTFATSDADNHTVMVDDQEVINKASNNKIKLEKGRLYQIK--- 117
DB 890 FPIYPMITEIKKCTLSFFCNSLENILM-----KXANFLKTLSPERNKYQISTPII 941
QY 118 -----IQYQRENTKGLDFKLWTDSONKK--EVISSDNLQLPK 157
```

Db 942 LLSITSYSSYGGKKXVLMKYEKTLKANTRIQLK-AWKNFSNFOQSFEMFIENNMI----- 996
Qy 158 QKSSNSRKRKSTAGPTVPDRDNDGIPDS-LEVEGYTVDVKNK-----RTFLS-----PWI 207
Db 997 -----DIPLOPIFDPIILNIKVKYKVKYFIGETNISLVPVLPWI 1037
Qy 208 SNIHEKGLTKYKSSPEKWTASTADPYDFKV-----TGRIDKVSPEARHPLVAAYPIV 262
Db 1038 KNIDE-----VLY-----YLOAHDDYSETINMKNDINTFYKN-----KXAAALVTSAIS 1082
Qy 263 HVDMENILSKVEDQSTQNTDSET-----RTISKSTSTSRHTSEVHGNAEVAHSFFD 315
Db 1083 LADCEDTSLKEELINKYENDDEAWKEIPLNLDQENQKEDKNKTSOCHGVNTVNDGYN 1142
Qy 316 IGGVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTARKLANIRIVNTGTAPI- 374
Db 1143 -NGAYEMGYMETYNIKNDN-----NNNNNNYNNNNNNNNNNNNNNNNNNNNYNYAAPT 1192
Qy 375 ---NVNLPFTSLVGLKNOTLATIKAD--ENOLSOILAPNNYPPSKLAPALNAQDASS 429
Db 1193 SYNRRVL-----QNDTRNNRYNHSNNMMNNNNNNNNNNNNNNNNNNNNNNNNNN 1244
Qy 430 TPIITWYN-----OFLELEKTK-----QLRLDTQVYGNIAATN--FENGRRVVDGTGNSW 477
Db 1245 KGNLTLPNNNNIHHFNKLSNNKFSYLSRIQKDTYNIKNYNSIYKLFDDGIPETIILSYN 1304
Qy 478 SEVLPOIQUETTARIEN-----GKDLNLVERRIAVERN----- 509
Db 1305 VANYPIKILTSKYLINLHIPRFLYVEGGKLN--IEKFIKNIRVSVVDGILENLYDDIL 1363
Qy 510 -PSDPLETTKPDMTL-----KEALKIA-----FGFNPNNG-----NLQYQGDKI 547
Db 1364 IPSLPLIKKNDISCDNNYNNENKIEKQKFCGFEQFPFVEIIGGQIKCFKIKYRNLES 1423
Qy 548 TEFDENFQOQSONI-----KNOLA-----ELNATNIYVLDKIKLNKNNILIR 592
Db 1424 ENNPUSLKDITNQIFRNKFRGNKIPLYLIRVYVLRGILYGINNEYTANPLYIFSLG 1483
Qy 593 DK-----RFRHYDRNNI-----AVGADESIVVKEAHEVINSSTEGLLNIDKIRKILSGYI 643
Db 1484 EKTSLNRAFKFSNINPEFGCLWESEAFPE--DEILTSIVSAEDNYDKQINDIYIG-- 1539
Qy 644 VEIETEGLEKEVIN--DRYDMLNITSLRQDGKTFIDFKK--YND--KLPLYISNPNYKV-- 696
Db 1540 ----STE-----INLFRDWSKWEKWHMKKPKIPVEYRPLYSNIXIKHPKVVSSNNYTMN 1590
Qy 697 ---NVVAV-----TKENTIINPSENGDTS 717
Db 1591 SWNNIFSPFDIPNLYMTYTSPTKGNNNNNNNNNNNNS 1627
PRELIMINARY; PRT; 2940 AA.
ID Q8IHP9
AC Q8IHP9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF11_0480.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]_TaxID=36329;
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=1236864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Valdiva A.B.,

RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mangall C.,
RA Venter J.C., Carucci D.J., Hofman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum";
RL Nature 419:498-511 (2002).
DR EMBL, AB014843; AAN36060.1; --
KW Hypothetical protein.
SQ SEQUENCE 2940 AA; 348040 MM; BF8748D8A051BAD2 CRC64;
Query Match 5.2%; Score 197.5; DB 5; Length 2940;
Best Local Similarity 18.7%; Pred. No. 0.92;
Matches 162; Conservative 138; Mismatches 322; Indels 243; Gaps 40;
Qy 34 VVTSSTGDLSPSSSELENIPSENCYFQSAIWSGFIKVKYSDEYTFATSAADHVTMWVDD 93
Db 627 ITHNNNNNDISLNHNNNN-----NYIFMNNYMHNNI-----NNNYCYCYNNNTNHVNNYNN 679
Qy 94 QEVINKASNSK-----IRLEKGRLYQIKIQYQRENPTKGLDFK-- 133
Db 680 IYIQNHNDQNNAPILQPINNHHLAHINDLCYHSEKNEYTKISKHOMNNINPQOSNGKNQ 739
Qy 134 -----LYWTDSONK-----KEVISDNQLPDLKQKSSNRKSTSGAPT-- 174
Db 740 NDISNNINKNDYYNLNEQKILCDKXSYIKCDIPQKCDNTQDDENSEQONQYITNPSG 799
Qy 175 -----VPRDNDGIP--DSLEVEGYTVDVKNKRTFLSPWISNIEHEKGL----- 216
Db 800 HYKIKEQNNYVQHIDPYEHDNTNEMINTQNT-----NINLDPYIMNNDNNVNLNQ 853
Qy 217 ----TKYKSSPE--KWTASDPDSD--FEKVTGRIDK--NVSPARHPLVAAPIVHVD 265
Db 854 NYFDETKQKKEEDIMITEANTHNNVHTYQNNSHNMKNMLNNNTNSKSTQDYDLYHNN 913
Qy 266 ME-----NITLSK--NEDQST-----QNTDSSTRAISKNTS--T 295
Db 914 MENFNTYNNKKEIKHINDTSSVNNSSINQQNLCKNNKNTNDYQNRINKDNSIPD 973
Qy 296 SRHTSEVHGNAEVAHSFPDGGVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTA 355
Db 974 SSINNELKNNIQMNDY--ASNTYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1024
Qy 356 DT--ARLAN--TRYVNTGTAPLYNVLP--TTSVLGKQNTLATIKADENQLS-----Q 403
Db 1025 NVTPQNSNNYIPIYD-----PHLNPEYKHTINNQLNKNKNNNNISMDNDVYDSNTISH 1080
Qy 404 ILAPNNYPSKNLAPIALNAQDASSTPTIMYNNQFLELEKTKQLRLDTQVYGNIAATN 463
Db 1081 ISVNDNEYNSNFTNDMYNNNN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNN 1132
Qy 464 FENGRRVVDGTGNSWSEVLPCIOETTA--RIIFNG--KDLNLVERRIAVERNPSDPLETTKPD 521
Db 1133 LYNSSNYNITQHMPSSNDVFKIOKHANNIIMNGHKEEHLIEKK-----SKEETNKYNE 1185
Qy 522 TLKEALKIAFGFNEPNGNLQYQKDIPEFDFNFDOQTSONIKNQLAEL-----NA 571
Db 1186 QVYRSI-----NQNTLILKENEIDENDINTLQ--NLNKKDMMNDNNIINKLNNI 1237
Qy 572 TNYITVLDKIKLNKAKNIIILKRRPHYDRANTAVGADESIVVKEAHEVINSSTEGLLNI 631
Db 1238 NNIYT-----PYQNILKNNIEQFLNN-----KEVITKHAYTHSSNEININVKNI 1284
Qy 632 DKDIRKILSGYIVEIEDTEGLEKEVINDRYDMLNITSLRQDGKTF-----ID- 677
Db 1285 DTQI-----NIRKNNNDIILNK--EQKDISTNNNQNTYNTITSSVKNEYSILDN 1334
Qy 678 ---FKYNDKPLIYINP-----NYKNVYAVTKENT-----IINPSENG 714
Db 1335 SPTDKYKKNLNLNLISSYNTNDHNNYHDDIDDKTKNDKNKYBEKHKDINIYNNINSNG 1394
Qy 715 DTSTNGI-----KKILIF 727

Db 1395 NKQNFINSYFDLNENEKKKINIF 1419

Search completed: May 3, 2004, 19:40:06
Job time : 37.6261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:26:27 ; Search time 47,5168 seconds
(without alignments)
4227.791 Million cell updates/sec

Title: US-09-848-909A-16
Perfect score: 3648
Sequence: 1 EVKQENRLNESSESSQGL.....TSTNGIKKILFSSKKGYEIG 711

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3648	100.0	711	5	AAM51498 Anthrax P
2	3642	99.8	711	5	AAM51497 Anthrax P
3	3635	99.6	711	5	AAM51496 Anthrax P
4	3629	99.5	711	5	AAM52113 Anthrax P
5	3626	99.4	735	5	AAM51495 Anthrax P
6	3620	99.2	735	5	AAM51493 Anthrax P
7	3614	99.1	735	5	AAM51485 Anthrax P
8	3613	99.0	735	5	AAM51491 Anthrax P
9	3613	99.0	735	5	AAM51492 Anthrax P
10	3609	98.9	735	5	AAM51480 Anthrax P
11	3608	98.9	735	5	AAM51487 Anthrax P
12	3608	98.9	735	5	AAM51500 Anthrax P
13	3608	98.9	735	5	AAM51489 Anthrax P
14	3607	98.9	735	2	AAR60179 Anthrax P
15	3607	98.9	735	5	AAM51483 Anthrax P
16	3607	98.9	735	5	AAM51488 Anthrax P
17	3607	98.9	735	5	AAM51494 Anthrax P
18	3607	98.9	735	5	AAM51499 Anthrax P
19	3607	98.9	736	3	AAY56959 Anthrax P
20	3607	98.9	763	3	AAY56960 B. anthra
21	3607	98.9	764	3	AAY56958 B. anthra
22	3607	98.9	764	4	AAB47306 Wild type
23	3607	98.9	857	7	AD65872 Bacillus
24	3606	98.8	735	5	AAM51484 Anthrax P
25	3606	98.8	735	5	AAM51486 Anthrax P

26	3604	98.8	735	5	AAE18289	Bacillus
27	3604	98.8	764	6	AAE35717	Bacillus
28	3592	98.5	764	6	ABP71693	B. anthra
29	3562.5	97.7	903	2	AAE60183	PA(1-725)
30	3430	94.0	719	2	AAE60193	Modified
31	2891	79.2	595	5	AAE18288	Bacillus
32	2752	75.4	569	3	AAE56961	B. anthra
33	2342	64.2	487	5	AAE18285	Bacillus
34	2027	55.6	426	5	AAE18287	Bacillus
35	1478	40.5	318	5	AAE18286	Bacillus
36	1338	36.7	258	5	AAE18284	Bacillus
37	1275	35.0	288	5	AAE50707	Bacillus
38	821.5	22.5	1052	4	AAE07903	C. botuli
39	820	22.5	1032	4	AAE07901	C. botuli
40	816	22.4	1092	4	AAE07900	C. botuli
41	807.5	22.1	1112	4	AAE07902	C. botuli
42	792	21.7	721	6	AAE35719	Clostridi
43	759.5	20.8	880	2	AAW60224	Bacillus
44	759	20.8	881	3	AAE59277	MIS toxin
45	759	20.8	884	2	AAE91239	B. cereus

ALIGNMENTS

RESULT 1
AAM51498
ID AAM51498 standard; protein; 711 AA.
XX AAM51498;
XX
XX
DT 01-FEB-2002 (first entry)
XX Anthrax PA K397D/D425K/F427A/D2L2 deletion mutant.
DE Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX Bacillus anthracis.
OS Synthetic.
Key Location/Qualifiers
FT Misc-difference 302..303
FT /note= "The mutant comprises a deletion amino acids 302-325 of the D2L2 loop of the wildtype protein"
FT Misc-difference 373
FT /note= "Wild-type Lys substituted by Asp"
FT Misc-difference 401
FT /note= "Wild-type Asp substituted by Lys"
FT Misc-difference 403
FT /note= "Wild-type Phe substituted by Ala"
WO200182798-A2.
08-NOV-2001.
04-MAY-2001; 2001WO-US014372.
04-MAY-2000; 2000US-0201800P.
(HARD) HARVARD COLLEGE.
Collier RJ, Sellman BR;
WPI; 2002-017725/02.
Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of Bacillus anthracis.
Claim 4; Page; 77pp; English.
The invention relates to antibacterial agents comprising mutant forms of

CC	pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants in the B moiety of the pore-forming binary A-B anthrax toxin, where the B moiety is anthrax protective antigen (PA) and using these mutants or combinations of them for protecting against Bacillus anthracis infections in humans, especially as vaccines. Note: The present sequence is not given in the specification but is derived from the Bacillus anthracis wild-type PA protein sequence shown in figure 13 (AAMS1483)	
XX	Sequence 711 AA;	
SY	Query Match 100.0%; Score 3648; DB 5; Length 711; Best Local Similarity 100.0%; Pred. No. 7.6e-241; Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 EVKQENRLNSESSESSQGLGYYFSDLNFQAPMVVTSSTTGLSPSSSELENIPSENQYF 60	
DB	1 EVKQENRLNSESSESSQGLGYYFSDLNFQAPMVVTSSTTGLSPSSSELENIPSENQYF 60	
QY	61 QSAIWSGFIKVKKSDSEYTFATSDADNHVTVMVDDQEVINKASNSKIRLEKGLYQIKIY 120	
DB	61 QSAIWSGFIKVKKSDSEYTFATSDADNHVTVMVDDQEVINKASNSKIRLEKGLYQIKIY 120	
QY	121 QRENPTXGLDFKLYWTDSONKKEVVISDNLQLPELKQKSSNRKCRSTSAGTVPDRDN 180	
DB	121 QRENPTXGLDFKLYWTDSONKKEVVISDNLQLPELKQKSSNRKCRSTSAGTVPDRDN 180	
QY	181 DGIPDSLEVGVTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASTADSEYSDFEKV 240	
DB	181 DGIPDSLEVGVTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASTADSEYSDFEKV 240	
QY	241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300	
DB	241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300	
QY	301 SNSNSTVAIDHSLSLAGERWAEITMGLNTADTARLANANTRYNTGTAPIYNVLPTTSIV 360	
DB	301 SNSNSTVAIDHSLSLAGERWAEITMGLNTADTARLANANTRYNTGTAPIYNVLPTTSIV 360	
QY	361 LGKNQPLATIKADENQLSQILAPNNYYPKKNLAPIALNAQKADASTPTIMNYNQFLELEK 420	
DB	361 LGKNQPLATIKADENQLSQILAPNNYYPKKNLAPIALNAQKADASTPTIMNYNQFLELEK 420	
QY	421 TKQLRLDTQVYGNATYTNFNGRVVDVDTGSNWSEVLPOIQTETARIIFNGKDLNIVER 480	
DB	421 TKQLRLDTQVYGNATYTNFNGRVVDVDTGSNWSEVLPOIQTETARIIFNGKDLNIVER 480	
QY	481 IAAVNSPDPLETTKPDWTLKEALKIAGFNPNPNGNLQYQKGKITEFDNFDQOTSQNIK 540	
DB	481 IAAVNSPDPLETTKPDWTLKEALKIAGFNPNPNGNLQYQKGKITEFDNFDQOTSQNIK 540	
QY	541 QLAELNATNIYVLDKIKLANAKNTLIEDKSPHYDRNNIAVGADESVVKEAHREVINSST 600	
DB	541 QLAELNATNIYVLDKIKLANAKNTLIEDKSPHYDRNNIAVGADESVVKEAHREVINSST 600	
QY	601 EGGLELNDIKDIRKILSGYIETDEGLKEVINDRYDMLNISSLRQDGKTFIDPKKYNK 660	
DB	601 EGGLELNDIKDIRKILSGYIETDEGLKEVINDRYDMLNISSLRQDGKTFIDPKKYNK 660	
QY	661 LPLYSINPNYKNNVAVTKNTIINPSENGDTSTNGIKKILLFSKKGYEIG 711	
DB	661 LPLYSINPNYKNNVAVTKNTIINPSENGDTSTNGIKKILLFSKKGYEIG 711	
RESULT 2		
ID	AAMS1497	
XX	AAVMS1497 standard; protein; 711 AA.	
AC	AAMS1497;	
XX		
DT	01-FEB-2002 (first entry)	
XX	Anthrax PA K397D/F427A/D212 deletion mutant.	
XX		

QY 361 LGKQTLATIKADENQLSQILAPNNYPSKVLAPIALNAQKADASTPTITWYNNQFLELEK 420
DB 361 LGKQTLATIKADENQLSQILAPNNYPSKVLAPIALNAQKADASTPTITWYNNQFLELEK 420
QY 421 TKQLRLDQVGNATYNFENGRVVDGTSNWSEVLPOIQETTARIIFNGKDLNLVERR 480
DB 421 TKQLRLDQVGNATYNFENGRVVDGTSNWSEVLPOIQETTARIIFNGKDLNLVERR 480
QY 481 IAAVNPSPDPLETTKPDMTLKEALKIAPGFNEPNGNLQYQKDIETFDNFDDQTSQNIKN 540
DB 481 IAAVNPSPDPLETTKPDMTLKEALKIAPGFNEPNGNLQYQKDIETFDNFDDQTSQNIKN 540
QY 541 QLAELNATNIYTVLDKIKLNAKNILIRDRKFRHYDRNNIAVGADESUVKEAHEVINSST 600
DB 541 QLAELNATNIYTVLDKIKLNAKNILIRDRKFRHYDRNNIAVGADESUVKEAHEVINSST 600
QY 601 EGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNISLRQDGTFFDFKKNYNDK 660
DB 601 EGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNISLRQDGTFFDFKKNYNDK 660
QY 661 LPLYISNPYKVNVAVTKENTIINPSENGDTSTNGIKKILIPSKGYEIG 711
DB 661 LPLYISNPYKVNVAVTKENTIINPSENGDTSTNGIKKILIPSKGYEIG 711

RESULT 3

AAM51496

ID AAM51496 standard; protein; 711 AA.

XX AC AAM51496;

XX DT 01-FEB-2002 (first entry)

XX DE Anthrax PA F427A/D2L2 deletion mutant.

XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX OS Bacillus anthracis.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 302..303

XX FT /note= "The mutant comprises a deletion amino acids 302-

XX FT 325 of the D2L2 loop of the wildtype protein"

XX FT Misc-difference 403

XX FT /note= "Wild-type Phe substituted by Ala"

XX PN WO200182788-A2.

XX PD 08-NOV-2001.

XX PF 04-MAY-2001; 2001WO-05014372.

XX PR 04-MAY-2000; 2000US-0201800P.

XX PA (HARD) HARVARD COLLEGE.

XX PI Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

XX FT Protecting humans against anthrax using mutant B groups (anthrax

XX FT protective antigens) of the pore-forming binary A-B toxin of Bacillus

XX FT anthracis.

XX PS Claim 4; Page; 77pp; English.

XX CC The invention relates to antibacterial agents comprising mutant forms of

XX CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants

XX CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B

XX CC moiety is anthrax protective antigen (PA) and using these mutants or

CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAM51483)

XX Sequence 711 AA;

QY Query Match 99.6%; Score 3635; DB 5; Length 711;

DB Best Local Similarity 99.7%; Pred. No. 5.9e-240;

QY Matches 709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSOGLLYFSDLNFOAPMVVTSTTGDLSIPSELENISENOYF 60

DB 1 EVKQENRLNSESSESSOGLLYFSDLNFOAPMVVTSTTGDLSIPSELENISENOYF 60

QY 61 QSAIWGFIKVKSDYEYFATSADNHVTWVDQEVINKASNNKIRLEKGLYQIKIY 120

DB 61 QSAIWGFIKVKSDYEYFATSADNHVTWVDQEVINKASNNKIRLEKGLYQIKIY 120

QY 121 QRENTEKGLDFKLTWYDSQNKKEVISSDNLQLPKQKSSNRKRSTAGTVPDNRN 180

DB 121 QRENTEKGLDFKLTWYDSQNKKEVISSDNLQLPKQKSSNRKRSTAGTVPDNRN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIEHKKGLTKYKSSPEKWSASTDPYSDFEKT 240

DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIEHKKGLTKYKSSPEKWSASTDPYSDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSTRITSKNTSRTHT 300

DB 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSTRITSKNTSRTHT 300

QY 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVLPPTSLV 360

DB 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVLPPTSLV 360

QY 361 LGKQTLATIKADENQLSQILAPNNYPSKVLAPIALNAQKADASTPTITWYNNQFLELEK 420

DB 361 LGKQTLATIKADENQLSQILAPNNYPSKVLAPIALNAQKADASTPTITWYNNQFLELEK 420

QY 421 TKQLRLDQVGNATYNFENGRVVDGTSNWSEVLPOIQETTARIIFNGKDLNLVERR 480

DB 421 TKQLRLDQVGNATYNFENGRVVDGTSNWSEVLPOIQETTARIIFNGKDLNLVERR 480

QY 481 IAAVNPSPDPLETTKPDMTLKEALKIAPGFNEPNGNLQYQKDIETFDNFDDQTSQNIKN 540

DB 481 IAAVNPSPDPLETTKPDMTLKEALKIAPGFNEPNGNLQYQKDIETFDNFDDQTSQNIKN 540

QY 541 QLAELNATNIYTVLDKIKLNAKNILIRDRKFRHYDRNNIAVGADESUVKEAHEVINSST 600

DB 541 QLAELNATNIYTVLDKIKLNAKNILIRDRKFRHYDRNNIAVGADESUVKEAHEVINSST 600

QY 601 EGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNISLRQDGTFFDFKKNYNDK 660

DB 601 EGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNISLRQDGTFFDFKKNYNDK 660

QY 661 LPLYISNPYKVNVAVTKENTIINPSENGDTSTNGIKKILIPSKGYEIG 711

DB 661 LPLYISNPYKVNVAVTKENTIINPSENGDTSTNGIKKILIPSKGYEIG 711

RESULT 4

AAM52113

ID AAM52113 standard; protein; 711 AA.

XX AC AAM52113;

XX DT 01-FEB-2002 (first entry)

XX DE Anthrax PA D2L2 deletion mutant.

XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX XX

Query Match 99.4%; Score 3626; DB 5; Length 735;
 Best Local Similarity 96.7%; Pred. No. 2.6e-239;
 Matches 711; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
 DB 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
 DB 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDQSNKKEVSSDNQLQPELKQKSSNRKRGSTAGTVPDRDN 180
 DB 121 QRENPTKGLDFKLYWTDQSNKKEVSSDNQLQPELKQKSSNRKRGSTAGTVPDRDN 180

QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKWSSTASDPYDFEKT 240
 DB 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKWSSTASDPYDFEKT 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
 DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
 DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 337 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPPSKNLAPIA 396
 DB 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPPSKNLAPIA 420

QY 397 LNAQKDASSPTITMNNYQFLEKTKQLRLDQDVYGNIAATYNFENGVRVDTGSNNSEV 456
 DB 421 LNAQKDASSPTITMNNYQFLEKTKQLRLDQDVYGNIAATYNFENGVRVDTGSNNSEV 480

QY 457 LPOIQTETAILFNGKDLNIVERRIAANVPSDPLETTKPDMTLKALKIAPFNPNGNL 516
 DB 481 LPOIQTETAILFNGKDLNIVERRIAANVPSDPLETTKPDMTLKALKIAPFNPNGNL 540

QY 517 QYQCKDITEDFDFDQTSQNIKNQLAELNATNIYVLDKIKLNAMKILIRDFHYDR 576
 DB 541 QYQCKDITEDFDFDQTSQNIKNQLAELNATNIYVLDKIKLNAMKILIRDFHYDR 600

QY 577 NNTAVGADESUVKAEHREVINSSTEGILLNIDKIRKILSGYIVIEDTEGLKEVINDRY 636
 DB 601 NNTAVGADESUVKAEHREVINSSTEGILLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660

QY 637 DMLNSSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 696
 DB 661 DMLNSSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720

QY 697 IKKILIPSKKGYBIG 711
 DB 721 IKKILIPSKKGYBIG 735

RESULT 6
 ID AAM51493 standard; protein; 735 AA.
 XX AAM51493;
 XX 01-FEB-2002 (first entry)
 XX Anthrax PA mutant K397D/D425K.
 XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
 XX Bacillus anthracis.
 OS Synthetic.

XX Key Location/Qualifiers
 FT Misc-difference /note "Wild-type Lys substituted by Asp"
 FT Misc-difference /note "Wild-type Asp substituted by Lys"
 XX W0200182788-A2.
 XX 08-NOV-2001.
 XX 04-MAY-2001; 2001WO-US014372.
 XX 04-MAY-2000; 2000US-0201800P.
 XX (HARD) HARVARD COLLEGE.
 XX Collier RJ, Sellman BR;
 XX WPI; 2002-017725/02.
 XX Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.
 XX Claim 4; Page; 77pp; English.
 XX The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections
 CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)

XX Sequence 735 AA;

Query Match 99.2%; Score 3620; DB 5; Length 735;
 Best Local Similarity 96.6%; Pred. No. 6.6e-239;
 Matches 710; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
 DB 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
 DB 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDQSNKKEVSSDNQLQPELKQKSSNRKRGSTAGTVPDRDN 180
 DB 121 QRENPTKGLDFKLYWTDQSNKKEVSSDNQLQPELKQKSSNRKRGSTAGTVPDRDN 180

QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKWSSTASDPYDFEKT 240
 DB 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKWSSTASDPYDFEKT 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
 DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
 DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 337 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPPSKNLAPIA 396
 DB 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPPSKNLAPIA 420

QY 397 LNAQKDASSPTITMNNYQFLEKTKQLRLDQDVYGNIAATYNFENGVRVDTGSNNSEV 456
 DB 421 LNAQKDASSPTITMNNYQFLEKTKQLRLDQDVYGNIAATYNFENGVRVDTGSNNSEV 480

QY 457 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWTLKEALKIAPGPNPGL 516
 DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWTLKEALKIAPGPNPGL 540
 QY 517 QYQCKDITEFDNPDQQTSONIKQLAELNATNIYTVLDKIKLNAGNILLIRDKRPFHYDR 576
 DB 541 QYQCKDITEFDNPDQQTSONIKQLAELNATNIYTVLDKIKLNAGNILLIRDKRPFHYDR 600
 QY 577 NNIAGADESVVKEAHEVINSSTEGLLINIDKIDIKILSGYIVLEDETEGLKEVINDRY 636
 DB 601 NNIAGADESVVKEAHEVINSSTEGLLINIDKIDIKILSGYIVLEDETEGLKEVINDRY 660
 QY 637 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPYKVNYYAVTKENTIINPSENGDTSTNG 696
 DB 661 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPYKVNYYAVTKENTIINPSENGDTSTNG 720
 QY 697 IKKILIFSKGYEIG 711
 DB 721 IKKILIFSKGYEIG 735

RESULT 7

AA51485
 ID AA51485 standard; protein; 735 AA.

AC AA51485;

DT 01-FEB-2002 (first entry)

DE Anthrax PA mutant K397D.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX Bacillus anthracis.

OS Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 397

FT /note= "Wild-type Lys substituted by Asp"

XX WO200182788-A2.

PD 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

XX (HARD) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

DR Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.

PS Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AA52113 and AA51484-AA51500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections
 CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AA51483)

XX Sequence 735 AA;

Query Match 99.1%; Score 3614; DB 5; Length 735;

Best Local Similarity 96.5%; Pred. No. 1.7e-236; Mismatches 2; Indels 24; Gaps 1;
 Matches 709; Conservative 0;

QY 1 EVKQENRLLNSESSESSQGLLYFFSDLNFOAPMVVTSITGDLSPSSSELENIPSENYCF 60
 DB 1 EVKQENRLLNSESSESSQGLLYFFSDLNFOAPMVVTSITGDLSPSSSELENIPSENYCF 60
 QY 61 QSAIWGSFTKVKKSDBYTFATSADNHNVTWVDDQVINKASNNKIRLEKGRLYQIKIY 120
 DB 61 QSAIWGSFTKVKKSDBYTFATSADNHNVTWVDDQVINKASNNKIRLEKGRLYQIKIY 120
 QY 121 QRENTEKGLDFKLXWTDSONKKEVISSDNLQLPKQKSNRSKRKSTAGTPVPPDRN 180
 DB 121 QRENTEKGLDFKLXWTDSONKKEVISSDNLQLPKQKSNRSKRKSTAGTPVPPDRN 180
 QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASPYSDFEKVT 240
 DB 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASPYSDFEKVT 240
 QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMEIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
 DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMEIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
 QY 301 -----SNSNSTVAIDHSISLAGERTWAETMGINTADTARL 336
 DB 301 SEVHGNAEVHSAFFDIGSVSAGFSNSNSTVAIDHSISLAGERTWAETMGINTADTARL 360
 QY 337 NANIRYVNTGTAPIYNNVLPTTSLVIGKQQTATIKADENQLSQILAPNNYPPSKNLAPIA 396
 DB 361 NANIRYVNTGTAPIYNNVLPTTSLVIGKQQTATIKADENQLSQILAPNNYPPSKNLAPIA 420
 QY 397 LNAQKDSATPTIMYNNQFLEKTKQLRLDQVYGNIAIYNFENGVRVDTGSNWSEV 456
 DB 421 LNAQKDSATPTIMYNNQFLEKTKQLRLDQVYGNIAIYNFENGVRVDTGSNWSEV 480
 QY 457 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWTLKEALKIAPGPNPGL 516
 DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWTLKEALKIAPGPNPGL 540
 QY 517 QYQCKDITEFDNPDQQTSONIKQLAELNATNIYTVLDKIKLNAGNILLIRDKRPFHYDR 576
 DB 541 QYQCKDITEFDNPDQQTSONIKQLAELNATNIYTVLDKIKLNAGNILLIRDKRPFHYDR 600
 QY 577 NNIAGADESVVKEAHEVINSSTEGLLINIDKIDIKILSGYIVLEDETEGLKEVINDRY 636
 DB 601 NNIAGADESVVKEAHEVINSSTEGLLINIDKIDIKILSGYIVLEDETEGLKEVINDRY 660
 QY 637 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPYKVNYYAVTKENTIINPSENGDTSTNG 696
 DB 661 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPYKVNYYAVTKENTIINPSENGDTSTNG 720
 QY 697 IKKILIFSKGYEIG 711
 DB 721 IKKILIFSKGYEIG 735

RESULT 8

AA51491
 ID AA51491 standard; protein; 735 AA.

AC AA51491;

XX 01-FEB-2002 (first entry)

XX Anthrax PA mutant D425K.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX Bacillus anthracis.

OS Synthetic.

XX

Key Location/Qualifiers
Misc-difference 425 /note= "Wild-type Asp substituted by Lys"
WO200182788-A2.
08-NOV-2001.
04-MAY-2001; 2001WO-US014372.
04-MAY-2000; 2000US-0201800P.
(HARD) HARVARD COLLEGE.
Collier RJ, Sellman BR;
WPI; 2002-017725/02.
Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of Bacillus anthracis.
Claim 4; Page; 77pp; English.
The invention relates to antibacterial agents comprising mutant forms of pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants in the B moiety of the pore-forming binary A-B anthrax toxin, where the B moiety is anthrax protective antigen (PA) and using these mutants or compositions of them for protecting against Bacillus anthracis infections in humans, especially as vaccines. Note: The present sequence is not given in the specification but is derived from the Bacillus anthracis wild-type PA protein sequence shown in figure 13 (AAMS1483)
Sequence 735 AA;
Query Match 99.0%; Score 3613; DB 5; Length 735;
Best Local Similarity 96.5%; Pred. No. 2e-238;
Matches 709; Conservative 0; Mismatches 2; Indels 24; Gaps 1;
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSELENIPSENQYF 60
QY 61 QSATWSGFIKVKSDYFATSAADNHVTWVDDDEVINKASNSKIRLEKRLYQIKIQY 120
DB 61 QSATWSGFIKVKSDYFATSAADNHVTWVDDDEVINKASNSKIRLEKRLYQIKIQY 120
QY 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKLPKOKSSNRKRGSTSGPTVPDRN 180
DB 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKLPKOKSSNRKRGSTSGPTVPDRN 180
QY 181 DGIPOSLVEGTVGVKNTKRTFLSPWISNIHEKGLTKYKSPKWSFASDPYDFEKT 240
DB 181 DGIPOSLVEGTVGVKNTKRTFLSPWISNIHEKGLTKYKSPKWSFASDPYDFEKT 240
QY 241 GRIDKNVSEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300
DB 241 GRIDKNVSEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAGTNGINTADTARL 336
DB 301 SEVHGNAEVSFFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTNGINTADTARL 360
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSOILAPNNYPSKNLAPIA 420
QY 397 LNAQKXASPTITMNYNPFLEKTKQLRLDPTDQVYGNIAATYVNFGRVVDVTGNSGEV 456
DB 421 LNAQKXASPTITMNYNPFLEKTKQLRLDPTDQVYGNIAATYVNFGRVVDVTGNSGEV 480
QY 457 LPQIOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKALKIAFGFNPNGNL 516

DB 481 LPQIOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKALKIAFGFNPNGNL 540
QY 517 QYQKDIITEFDNFQDQTSQNIKNQALAEINATNIYTVLDIKLNAKNILIRDKRFHYDR 576
DB 541 QYQKDIITEFDNFQDQTSQNIKNQALAEINATNIYTVLDIKLNAKNILIRDKRFHYDR 600
QY 577 NNIAVGADESUVKAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 636
DB 601 NNIAVGADESUVKAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
QY 637 DMLNISSLRQDKTFIDFKYNDKLPYISNPYKVVAVTKENTINPSENGDTSTNG 696
DB 661 DMLNISSLRQDKTFIDFKYNDKLPYISNPYKVVAVTKENTINPSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
DB 721 IKKILIFSKKGYEIG 735
RESULT 9
AAMS1492
ID AAMS1492 standard; protein; 735 AA.
XX AAMS1492;
AC AAMS1492;
XX 01-FEB-2002 (first entry)
DT Anthrax PA mutant F427A.
DE Anthrax PA mutant F427A.
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.
XX Bacillus anthracis.
OS Synthetic.
XX Key Location/Qualifiers
FH Key Location/Qualifiers
FT Misc-difference 427 /note= "Wild-type Phe substituted by Ala"
XX WO200182788-A2.
PN 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US014372.
XX 04-MAY-2000; 2000US-0201800P.
PR (HARD) HARVARD COLLEGE.
PA Collier RJ, Sellman BR;
PI WPI; 2002-017725/02.
XX Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of Bacillus anthracis.
XX Claim 4; Page; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants in the B moiety of the pore-forming binary A-B anthrax toxin, where the B moiety is anthrax protective antigen (PA) and using these mutants or compositions of them for protecting against Bacillus anthracis infections in humans, especially as vaccines. Note: The present sequence is not given in the specification but is derived from the Bacillus anthracis wild-type PA protein sequence shown in figure 13 (AAMS1483)
XX Sequence 735 AA;
SQ Query Match 99.0%; Score 3613; DB 5; Length 735;
Best Local Similarity 96.5%; Pred. No. 2e-238;
Matches 709; Conservative 0; Mismatches 2; Indels 24; Gaps 1;

QY 1 EVKQENLLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
DB 1 EVKQENLLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDYFTFATSADNHNVTMVDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDYFTFATSADNHNVTMVDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTAGTPVDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTAGTPVDRDN 180
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSPEKVT 240
DB 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSPEKVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 336
DB 301 SEVHGNAEYHAFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 397 LNAQKDASSPTITMNNYQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTQSNWSEV 456
DB 421 LNAQKDASSPTITMNNYQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTQSNWSEV 480
QY 457 LPOIQTETARIIFNGKDLNVERRIAAYVNPSPDPLETTKPDMLKEALKIAGFNPENGNL 516
DB 481 LPOIQTETARIIFNGKDLNVERRIAAYVNPSPDPLETTKPDMLKEALKIAGFNPENGNL 540
QY 517 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 576
DB 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 600
QY 577 NNTIAGDAESVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEDETEGLKEVINDRY 636
DB 601 NNTIAGDAESVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEDETEGLKEVINDRY 660
QY 637 DMLNISSLRQDKTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 696
DB 661 DMLNISSLRQDKTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 697 IKKILIFSKGYEIG 711
DB 721 IKKILIFSKGYEIG 735

RESULT 10
AAM51490
ID AAM51490 standard; protein; 735 AA.
XX
AC AAM51490;
XX
DT 01-FEB-2002 (first entry)
DE Anthrax PA mutant D425E.
XX
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
OS Bacillus anthracis.
XX
FH Key Location/Qualifiers
FT Misc-difference 425 /note= "Wild-type Asp substituted by Glu"

XX WO200182788-A2.
PN
XX
PD 08-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014372.
XX
PR 04-MAY-2000; 2000US-0201800P.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Collier RJ, Sellman BR;
XX
XX WPI; 2002-017725/02.
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX
XX Claim 4; Page; 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
SQ Sequence 735 AA;
XX
Query Match 98.9%; Score 3609; DB 5; Length 735;
Best Local Similarity 96.3%; Pred. No. 3.7e-238;
Matches 708; Conservative 1; Mismatches 2; Indels 24; Gaps 1;
QY 1 EVKQENLLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
DB 1 EVKQENLLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDYFTFATSADNHNVTMVDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDYFTFATSADNHNVTMVDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTAGTPVDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTAGTPVDRDN 180
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSPEKVT 240
DB 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSPEKVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 336
DB 301 SEVHGNAEYHAFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 397 LNAQKDASSPTITMNNYQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTQSNWSEV 456
DB 421 LNAQKDASSPTITMNNYQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTQSNWSEV 480
QY 457 LPOIQTETARIIFNGKDLNVERRIAAYVNPSPDPLETTKPDMLKEALKIAGFNPENGNL 516
DB 481 LPOIQTETARIIFNGKDLNVERRIAAYVNPSPDPLETTKPDMLKEALKIAGFNPENGNL 540
QY 517 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 576

Db 541 QYQKDIETEDFNQDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNMLIRDRPHYDR 600
QY 577 NNTAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDTEGLKEVINDRY 636
Db 601 NNTAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDTEGLKEVINDRY 660
QY 637 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVYAVTKENTIINPSENGDTSTNG 696
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVYAVTKENTIINPSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735
RESULT 11
AAM51487 standard; protein; 735 AA.
XX AC AAM51487;
XX DT 01-FEB-2002 (first entry)
XX DE Anthrax PA mutant K397Q.
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX OS Bacillus anthracis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 397 /note= "Wild-type Lys substituted by Gln"
XX FT WO200182788-A2.
XX PN 08-NOV-2001.
XX PD 04-MAY-2001; 2001WO-US014372.
XX PF 04-MAY-2000; 2000US-0201800P.
XX PR (HARD) HARVARD COLLEGE.
XX PA Collier RJ, Sellman BR;
XX PI WPI; 2002-017725/02.
XX DR Protecting humans against anthrax using mutant B groups (anthrax
XX PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX PT anthracis.
XX PS Claim 4; Page; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAM51483)
XX SQ Sequence 735 AA;
Query Match 98.9%; Score 3608; DB 5; Length 735;
Best Local Similarity 96.3%; Pred. No. 4.4e-238;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYQF 60

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYQF 60
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHYTMWDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHYTMWDDQEVINKASNSNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNIQLPELKQKSSNSKRGKSTAGTPVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNIQLPELKQKSSNSKRGKSTAGTPVPDRN 180
QY 181 DGIPLDSLEVEGYTDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
Db 181 DGIPLDSLEVEGYTDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
QY 241 GRIDKNVSPKARHPLAAYPIVHVDMENIILSKNEDQSTQNTDGETRISKNTSRTHT 300
Db 241 GRIDKNVSPKARHPLAAYPIVHVDMENIILSKNEDQSTQNTDGETRISKNTSRTHT 300
QY 301 -----SNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYNNVLPSTSLVLGKNOTLATIKADENOLSOILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYNNVLPSTSLVLGKNOTLATIKADENOLSOILAPNNYPSKNLAPIA 420
QY 397 LNAQKDASPTITMNYNOFLELEKTKQLRLDQDQYGNIAFYNFENGRVVDVTSNWSSEV 456
Db 421 LNAQDFFSTPTITMNYNOFLELEKTKQLRLDQDQYGNIAFYNFENGRVVDVTSNWSSEV 480
QY 457 LPOIQTETARIIPNGKDLNVERRIAANVPSPLETTKPDMLKEALKIAGFNEPNGNL 516
Db 481 LPOIQTETARIIPNGKDLNVERRIAANVPSPLETTKPDMLKEALKIAGFNEPNGNL 540
QY 517 QYQKDIETEDFNQDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNMLIRDRPHYDR 576
Db 541 QYQKDIETEDFNQDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNMLIRDRPHYDR 600
QY 577 NNTAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDTEGLKEVINDRY 636
Db 601 NNTAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDTEGLKEVINDRY 660
QY 637 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVYAVTKENTIINPSENGDTSTNG 696
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVYAVTKENTIINPSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735
RESULT 12
AAM51500
XX ID AAM51500 standard; protein; 735 AA.
XX AC AAM51500;
XX DT 01-FEB-2002 (first entry)
XX DE Anthrax PA F427K mutant.
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX OS Bacillus anthracis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 427 /note= "Wild-type Phe substituted by Lys"
XX FT WO200182788-A2.
XX PN


```

Db 61 QSAIWSGFIKVKSDYFTFATSADNHTVMWVDOEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKSRSTAGPTVDDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKSRSTAGPTVDDRDN 180
QY 181 DGIPLSLEVGTYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKMWSTASDPYDFEYV 240
Db 181 DGIPLSLEVGTYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKMWSTASDPYDFEYV 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYNNVLTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 397 LNAQKASSTPTITMNYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 456
Db 421 LNAQKASSTPTITMNYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
QY 457 LPOIQUETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLKALKIARFNEPENG 516
Db 481 LPOIQUETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLKALKIARFNEPENG 540
QY 517 QYQKDIETFDNFDOQTSQNIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRFHYDR 576
Db 541 QYQKDIETFDNFDOQTSQNIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRFHYDR 600
QY 577 NNIAVGADESUVKEAHRVINSSTEGLLNIDKDIRKILSGYVIEDTEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAHRVINSSTEGLLNIDKDIRKILSGYVIEDTEGLKEVINDRY 660
QY 637 DMLNISLRQDGTFFDKKYNKDLFLYISNPNKYVNVYAVTKENTIINPSENGPTSTNG 696
Db 661 DMLNISLRQDGTFFDKKYNKDLFLYISNPNKYVNVYAVTKENTIINPSENGPTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735

```

RESULT 14

AA60179
ID AAR60179 standard; protein; 735 AA.

AC AAR60179;
XX

XX 25-MAR-2003 (revised)
DT 03-APR-1995 (first entry)

XX

XX Protective antigen of *Bacillus anthracis*.

XX Anthrax; *Bacillus anthracis*; fusion protein; protective antigen;
XX protective antigen; cell killing; targeting; toxin;
XX intracellular; HIV; human immunodeficiency virus; toxin.

XX *Bacillus anthracis*.

XX WO9418332-A2.

XX 18-AUG-1994.

XX 14-FEB-1994; 94WO-US001624.

XX 12-FEB-1993; 93US-00021601.

XX 25-JUN-1993; 93US-00082849.

```

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Leppla SH, Klimpel K, Arora N, Singh Y, Nichols PJ;
XX DR WPI; 1994-279753/34.
XX DR N-PSDB; AAQ70180.
XX PT Nucleic acid encoding anthrax toxin fusion protein - useful for
XX PT targeting toxin to specific cells, eg for killing tumour cells or HIV-
XX PS infected cells.
XX PS Disclosure; Page 81-83; 124pp; English.
XX CC The sequence encoding the protective antigen of Bacillus anthracis may be
XX CC used in the construction of a nucleic acid which encodes a fusion protein
XX CC comprising the anthrax protective antigen binding domain of the native
XX CC anthrax lethal factor and a sequence encoding an activity inducing domain
XX CC of a second protein. The fusion proteins are useful for the specific
XX CC killing of tumour cells or the killing of cells infected with
XX CC intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to
XX CC correct PN field.)
XX SQ Sequence 735 AA;

```

Query Match 98.9%; Score 3607; DB 2; Length 735;

Best Local Similarity 96.3%; Pred. No. 5,1e-238;

Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

```

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSTTGDLSIPSELENISENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSTTGDLSIPSELENISENQYF 60
QY 61 QSAIWSGFIKVKSDYFTFATSADNHTVMWVDOEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYFTFATSADNHTVMWVDOEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKSRSTAGPTVDDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKSRSTAGPTVDDRDN 180
QY 181 DGIPLSLEVGTYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKMWSTASDPYDFEYV 240
Db 181 DGIPLSLEVGTYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKMWSTASDPYDFEYV 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYNNVLTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 397 LNAQKASSTPTITMNYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 456
Db 421 LNAQKASSTPTITMNYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
QY 457 LPOIQUETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLKALKIARFNEPENG 516
Db 481 LPOIQUETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLKALKIARFNEPENG 540
QY 517 QYQKDIETFDNFDOQTSQNIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRFHYDR 576
Db 541 QYQKDIETFDNFDOQTSQNIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRFHYDR 600
QY 577 NNIAVGADESUVKEAHRVINSSTEGLLNIDKDIRKILSGYVIEDTEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAHRVINSSTEGLLNIDKDIRKILSGYVIEDTEGLKEVINDRY 660

```

QY 637 DMLNLSLRQDQKTFIDFKKNDKLPYISNPYKVVAVTKENTIIINSENGDTSTNG 696
DB 661 DMLNLSLRQDQKTFIDFKKNDKLPYISNPYKVVAVTKENTIIINSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
DB 721 IKKILIFSKKGYEIG 735

RESULT 15
ID AAM51483
XX AAM51483 standard; protein; 735 AA.
XX AC AAM51483;
XX DT 01-FEB-2002 (first entry)
XX DE Anthrax PA protein.
XX EW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.
XX OS Bacillus anthracis.
XX PN WO200182788-A2.
XX PD 08-NOV-2001.
XX PF 04-MAY-2001; 2001WO-US014372.
XX PR 04-MAY-2000; 2000US-0201800P.
XX PA (HARD) HARVARD COLLEGE.
XX PI Collier RJ, Sellman BR;
XX XX WPI; 2002-017725/02.
XX DR N-PSDB; AA199904.
XX PT Protecting humans against anthrax using mutant B groups (anthrax
XX PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX PT anthracis.
XX PS Disclosure; Fig 13; 77pp; English.
XX CC The invention relates to antibacterial agents comprising mutant forms of
XX CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX CC moiety is anthrax protective antigen (PA) and using these mutants or
XX CC compositions of them for protecting against Bacillus anthracis infections
XX CC in humans, especially as vaccines. The present sequence is that of the
XX CC anthrax PA protein
XX SQ Sequence 735 AA;

Query Match 98.9%; Score 3607; DB 5; Length 735;
Best Local Similarity 96.3%; Pred. No. 5.1e-238;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLLNESSSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSLENIIPSENQYF 60
DB 1 EVKQENRLLNESSSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSLENIIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEVTFATSDADNVTWVDQEVINKASNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEVTFATSDADNVTWVDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRRSTAGTVPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRRSTAGTVPDRN 180
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEYKT 240

DB 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEYKT 240
QY 241 GRIDKNVSPFARHPLVAAAPVHVVDMENIILSKNEDQSTQNTDSRTRISKNTSRTHT 300
DB 241 GRIDKNVSPFARHPLVAAAPVHVVDMENIILSKNEDQSTQNTDSRTRISKNTSRTHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 336
DB 301 SEVHGNAEVHASFDFDIGGSVAGPFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
QY 337 NANIRYVNTGTAPIYVNLPTTSLVIGKNOTLATIKADENQLSQILAPNNYVPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYVNLPTTSLVIGKNOTLATIKADENQLSQILAPNNYVPSKNLAPIA 420
QY 397 LNAQKDASSTPITMNYNOFLELEKTKQLRLDQVYGNIAVYNFENGVRVDTGNSWSEV 456
DB 421 LNAQDDFSSTPITMNYNOFLELEKTKQLRLDQVYGNIAVYNFENGVRVDTGNSWSEV 480
QY 457 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 516
DB 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540
QY 517 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILIRDKFHYDR 576
DB 541 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILIRDKFHYDR 600
QY 577 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIDRKILSGYVIEIDTEGLKEVINDRY 636
DB 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIDRKILSGYVIEIDTEGLKEVINDRY 660
QY 637 DMLNLSLRQDQKTFIDFKKNDKLPYISNPYKVVAVTKENTIIINSENGDTSTNG 696
DB 661 DMLNLSLRQDQKTFIDFKKNDKLPYISNPYKVVAVTKENTIIINSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
DB 721 IKKILIFSKKGYEIG 735

Search completed: May 3, 2004, 19:36:05
Job time : 49.5168 secs

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082.849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-4

Query Match 98.9%; Score 3607; DB 1; Length 735;
Best Local Similarity 96.3%; Pred. No. 6.2e-261;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

1 EVKQENRLNSESQGLLYFSDLNFPQPMVVTSTTGDLSIPSSLENIIPSENQYF 60
61 QSAIWSGFIKVKSDSEYTFATSAHNVTWVDDQEVINKASNNKIRLEKGRLYQIKIQY 120
61 QSAIWSGFIKVKSDSEYTFATSAHNVTWVDDQEVINKASNNKIRLEKGRLYQIKIQY 120
121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSNRSKRKSTAGPTVPDRN 180
121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSNRSKRKSTAGPTVPDRN 180
181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKSTASDPYDFPKVT 240
181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKSTASDPYDFPKVT 240
241 GRIDKNVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
241 GRIDKNVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
301 SEVHGNAEVAHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
337 NANIRYVNTGTAPIYVLPPTSLVLGKNTLATIKADENQLSQILAPNNYPSKNLAPIA 396
361 NANIRYVNTGTAPIYVLPPTSLVLGKNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
397 LNAQKDAASSTPITMNYNQFLELEKTKQLRLDTPQVYGNIAATYFNGRVRVDTGSNWSEV 456
421 LNAQKDAASSTPITMNYNQFLELEKTKQLRLDTPQVYGNIAATYFNGRVRVDTGSNWSEV 480
457 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 516
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
517 QYQKDIITEFDNFDOOTSONIKNQLAELNATNIYVLDKIKNAKNNILIRDKRFHYDR 576
541 QYQKDIITEFDNFDOOTSONIKNQLAELNATNIYVLDKIKNAKNNILIRDKRFHYDR 600
577 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 636
601 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

RESULT 2
US-08-082-849B-4
Sequence 4, Application US/08082849B
Patent No. 5677274
GENERAL INFORMATION:
APPLICANT: Leppla, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
TITLE OF INVENTION: Related Methods
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082.849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-4

Query Match 98.9%; Score 3607; DB 1; Length 735;
Best Local Similarity 96.3%; Pred. No. 6.2e-261;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

1 EVKQENRLNSESQGLLYFSDLNFPQPMVVTSTTGDLSIPSSLENIIPSENQYF 60
1 EVKQENRLNSESQGLLYFSDLNFPQPMVVTSTTGDLSIPSSLENIIPSENQYF 60
61 QSAIWSGFIKVKSDSEYTFATSAHNVTWVDDQEVINKASNNKIRLEKGRLYQIKIQY 120
61 QSAIWSGFIKVKSDSEYTFATSAHNVTWVDDQEVINKASNNKIRLEKGRLYQIKIQY 120
121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSNRSKRKSTAGPTVPDRN 180
121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSNRSKRKSTAGPTVPDRN 180
181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKSTASDPYDFPKVT 240
181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKSTASDPYDFPKVT 240
241 GRIDKNVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
241 GRIDKNVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
301 SEVHGNAEVAHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
337 NANIRYVNTGTAPIYVLPPTSLVLGKNTLATIKADENQLSQILAPNNYPSKNLAPIA 396
361 NANIRYVNTGTAPIYVLPPTSLVLGKNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
397 LNAQKDAASSTPITMNYNQFLELEKTKQLRLDTPQVYGNIAATYFNGRVRVDTGSNWSEV 456
421 LNAQKDAASSTPITMNYNQFLELEKTKQLRLDTPQVYGNIAATYFNGRVRVDTGSNWSEV 480
457 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 516
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
517 QYQKDIITEFDNFDOOTSONIKNQLAELNATNIYVLDKIKNAKNNILIRDKRFHYDR 576
541 QYQKDIITEFDNFDOOTSONIKNQLAELNATNIYVLDKIKNAKNNILIRDKRFHYDR 600
577 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 636
601 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

QY 637 DMLNSSLRODGTFFIDFKYNDKLPYISNPNYKVVAVTKENTIIINPSENGDTSTNG 696
DB 661 DMLNSSLRODGTFFIDFKYNDKLPYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720
QY 697 IKKILIPSKGYEIG 711
DB 721 IKKILIPSKGYEIG 735
RESULT 3
PCT-US94-01624-4
Sequence 4, Application PC/TUS9401624
GENERAL INFORMATION:
APPLICANT: Leppa, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arota, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
RELATED METHODS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
STREET: Stewart Street Tower, 20th Floor, One Market
STREET: Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01624
FILING DATE: June 25, 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-4
Query Match 98.9%; Score 3607; DB 5; Length 735;
Best Local Similarity 96.3%; Pred. No. 6.2e-261;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;
QY 1 EVKQENRLNESSSQGLGYFFDLNFAQPMVVTSTTGLSPSSSELENIPSENYF 60
DB 1 EVKQENRLNESSSQGLGYFFDLNFAQPMVVTSTTGLSPSSSELENIPSENYF 60
QY 61 QSAIWSGFIKVKGSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKGSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLWYTDSONKKEVITSSDNLQLPELKQKSSNRKCRSTTSAGTVPDRDN 180
DB 121 QRENPTKGLDFKLWYTDSONKKEVITSSDNLQLPELKQKSSNRKCRSTTSAGTVPDRDN 180
QY 181 DGIIPDSLEVGTYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTADDPYSDPEKVT 240
DB 181 DGIIPDSLEVGTYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTADDPYSDPEKVT 240

QY 241 GRIDKNVSPEARHPVAAYPIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPEARHPVAAYPIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 336
DB 301 SEVHCNAEVSASFDPDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQSLQILAPNNYPPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQSLQILAPNNYPPSKNLAPIA 420
QY 397 LNAQKQDASSTPTWYNYNOFLELEKTKQLRLDQVYGNIAFYNFENGVRVYDTGSNWSEV 456
DB 421 LNAQDQFSSTPTWYNYNOFLELEKTKQLRLDQVYGNIAFYNFENGVRVYDTGSNWSEV 480
QY 457 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDPLETTKPDMTLKEALKIAGFNEPNGNL 516
DB 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDPLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 517 QYQGGKIDTEFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNAKMMILIRDKRFHYDR 576
DB 541 QYQGGKIDTEFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNAKMMILIRDKRFHYDR 600
QY 577 NNIAGVDESIVVKEAHREVINSSTGLLLNIDKIDRKILSGYIVIEDETEGLKEVINDRY 636
DB 601 NNIAGVDESIVVKEAHREVINSSTGLLLNIDKIDRKILSGYIVIEDETEGLKEVINDRY 660
QY 637 DMLNSSLRODGTFFIDFKYNDKLPYISNPNYKVVAVTKENTIIINPSENGDTSTNG 696
DB 661 DMLNSSLRODGTFFIDFKYNDKLPYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720
QY 697 IKKILIPSKGYEIG 711
DB 721 IKKILIPSKGYEIG 735
RESULT 4
US-08-021-601-12
Sequence 12, Application US/08021601
Patent No. 5591631
GENERAL INFORMATION:
APPLICANT: Leppa, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Nichols, Peter J.
APPLICANT: Arota, Naveen
APPLICANT: Singh, Yogendra
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
RELATED METHODS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Needle & Rosenberg, P.C.
STREET: 133 Carnegie Way, Suite 400
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021.601
FILING DATE: 19930212
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-021-601-12

Query Match 97.7%; Score 3562.5; DB 1; Length 903;
Best Local Similarity 95.5%; Pred. No. 1.8e-257;
Matches 701; Conservative 1; Mismatches 5; Indels 27; Gaps 2;
Qy 1 EVKQENLLNESSSQGLLYGYSDFLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENLLNESSSQGLLYGYSDFLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDGYTPATSAADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDGYTPATSAADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKSTAGPTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKSTAGPTVPDRN 180
Qy 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240
Db 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAGTMTGLNTADTARL 336
Db 301 SEVHGNAEVEHAFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTMTGLNTADTARL 360
Qy 337 NANIRYNTGTAPIYVLTPTSLVLGKNTLATIKADENQLSQILLAPNNYPSKNLAPIA 396
Db 361 NANIRYNTGTAPIYVLTPTSLVLGKNTLATIKADENQLSQILLAPNNYPSKNLAPIA 420
Qy 397 LNAQKQDASPTITMNVNQFLEKTKQLRLDTPQVYGNIAIYFNFGVRVDTGNSNWEV 456
Db 421 LNAQKQDASPTITMNVNQFLEKTKQLRLDTPQVYGNIAIYFNFGVRVDTGNSNWEV 480
Qy 457 LQIQETTARIIFNGKDLNLEVRRAAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 516
Db 481 LQIQETTARIIFNGKDLNLEVRRAAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540
Qy 517 QYQKDIITFDENFDQOTSONIKNOLAEALNATNIYVLDKIKINAKMILIRDKRFHYDR 576
Db 541 QYQKDIITFDENFDQOTSONIKNOLAEALNATNIYVLDKIKINAKMILIRDKRFHYDR 600
Qy 577 NNIAVGADESVMKAEHREVINSGTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 636
Db 601 NNIAVGADESVMKAEHREVINSGTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
Qy 637 DMLNISLRQDGKTFIDFKKYNKPLIYISNPNYKVVAVYKNTIINPSNGDTSTNG 696
Db 661 DMLNISLRQDGKTFIDFKKYNKPLIYISNPNYKVVAVYKNTIINPSNGDTSTNG 720
Qy 697 IKKIL---IFSKKG 707
Db 721 IKKILKVVLGKKG 734

RESULT 5
US-08-082-849B-12
Sequence 12, Application US/08082849B
Patent No. 5677274
GENERAL INFORMATION:
APPLICANT: Leppela, Stephen H.
APPLICANT: Klumpel, Kurt R.
APPLICANT: Arora, Naveen

APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
TITLE OF INVENTION: Related Methods
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-12

Query Match 97.7%; Score 3562.5; DB 1; Length 903;
Best Local Similarity 95.5%; Pred. No. 1.8e-257;
Matches 701; Conservative 1; Mismatches 5; Indels 27; Gaps 2;
Qy 1 EVKQENLLNESSSQGLLYGYSDFLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENLLNESSSQGLLYGYSDFLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDGYTPATSAADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDGYTPATSAADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKSTAGPTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKSTAGPTVPDRN 180
Qy 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240
Db 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAGTMTGLNTADTARL 336
Db 301 SEVHGNAEVEHAFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTMTGLNTADTARL 360
Qy 337 NANIRYNTGTAPIYVLTPTSLVLGKNTLATIKADENQLSQILLAPNNYPSKNLAPIA 396
Db 361 NANIRYNTGTAPIYVLTPTSLVLGKNTLATIKADENQLSQILLAPNNYPSKNLAPIA 420
Qy 397 LNAQKQDASPTITMNVNQFLEKTKQLRLDTPQVYGNIAIYFNFGVRVDTGNSNWEV 456

421 LNAQDFFSTPTIMYNOFLLEKTKQLRLDQVYGNIAATYFENGVRVDTGSGNWSEV 480
457 LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPENG 516
481 LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPENG 540
517 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAGNIIIRDKRPHYDR 576
541 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAGNIIIRDKRPHYDR 600
577 NNIAGDAESVVKAEHREVINSSTEGLLNIDKDIRKILSGYIIEIDTGLKEVINDRY 636
601 NNIAGDAESVVKAEHREVINSSTEGLLNIDKDIRKILSGYIIEIDTGLKEVINDRY 660
637 DMLNISSLRQDGTFFIDFKKNDKLPYISNPNKVNVAATKENTIIINSENGDTSTNG 696
661 DMLNISSLRQDGTFFIDFKKNDKLPYISNPNKVNVAATKENTIIINSENGDTSTNG 720
697 IKKIL---IFSKEG 707
721 IKKILKVVGLGKG 734
RESULT 6
PCT-US94-01624-12
; Sequence 12, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01624-12
Query Match 97.7%; Score 3562.5; DB 5; Length 903;
Best Local Similarity 95.5%; Pred. No. 1.8e-257;
Matches 701; Conservative 1; Mismatches 5; Indels 27; Gaps 2;
1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60

1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
61 QSAIWSGFIKVKSDIETFTATSDNHVMTWDDQVINKASNSKIRLEKGLYQIKQY 120
61 QSAIWSGFIKVKSDIETFTATSDNHVMTWDDQVINKASNSKIRLEKGLYQIKQY 120
121 QRENTEKGLDFKLWYTDSONKKEVSSDNLQQLPELKOKSSNSRKRKSTSGPTVPDRDN 180
121 QRENTEKGLDFKLWYTDSONKKEVSSDNLQQLPELKOKSSNSRKRKSTSGPTVPDRDN 180
181 DGIPDSLEVEGYTVDVQKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYSDFEKT 240
181 DGIPDSLEVEGYTVDVQKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYSDFEKT 240
241 GRIDKNVSEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSTRIISKNTSSTRTHT 300
241 GRIDKNVSEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSTRIISKNTSSTRTHT 300
301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
337 NANIRYVNTGTAPIYVNLPTTSVLGKQOTLATIKADENOLSOILAPNNYFESKNLAPIA 396
361 NANIRYVNTGTAPIYVNLPTTSVLGKQOTLATIKAKENOLSOILAPNNYFESKNLAPIA 420
397 LNAQKASSTPTIMYNOFLLEKTKQLRLDQVYGNIAATYFENGVRVDTGSGNWSEV 456
421 LNAQDFFSTPTIMYNOFLLEKTKQLRLDQVYGNIAATYFENGVRVDTGSGNWSEV 480
457 LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPENG 516
481 LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPENG 540
517 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAGNIIIRDKRPHYDR 576
541 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAGNIIIRDKRPHYDR 600
577 NNIAGDAESVVKAEHREVINSSTEGLLNIDKDIRKILSGYIIEIDTGLKEVINDRY 636
601 NNIAGDAESVVKAEHREVINSSTEGLLNIDKDIRKILSGYIIEIDTGLKEVINDRY 660
637 DMLNISSLRQDGTFFIDFKKNDKLPYISNPNKVNVAATKENTIIINSENGDTSTNG 696
661 DMLNISSLRQDGTFFIDFKKNDKLPYISNPNKVNVAATKENTIIINSENGDTSTNG 720
697 IKKIL---IFSKEG 707
721 IKKILKVVGLGKG 734
RESULT 7
US-08-082-849B-31
; Sequence 31, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082.849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021.601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-31

Query Match 94.0%; Score 3430; DB 1; Length 719;
Best Local Similarity 91.9%; Pred. No. 9.8e-248;
Matches 679; Conservative 4; Mismatches 28; Indels 48; Gaps 3;

QY 1 EVKQENLLNESSSQGLLYGFFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60
DB 1 EVKQENLLNESSSQGLLYGFFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60
QY 61 QSAIWSGFIKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNS-----RKKSTSGPTVP 176
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSNTATIMQKGNFLQSGTVP 180
QY 177 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSD 236
DB 181 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSD 240
QY 237 EKVTRIDKNSPEARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTS 296
DB 241 EKVTRIDKNSPEARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTS 300
QY 297 RTHT-----SNSNSTVAIDHSLSLAGERWAEWTMGINTAD 332
DB 301 RTHTSEVHGNAEVHAFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTMGINTAD 360
QY 333 TAILNANIRVNTGTAPIYVNLPTTSLVIGKQOTLATIKADENQLSQILAPNYPKSL 392
DB 361 TAILNANIRVNTGTAPIYVNLPTTSLVIGKQOTLATIKADENQLSQILAPNYPKSL 420
QY 393 APIALNAQDASPTITMNVNQFLEKTKQLRLDQVVGNIATVNFENGVRVDTGSN 452
DB 421 APIALNAQDDFSSTPTIMN-----YGNIAVNFENGVRVDTGSN 460
QY 453 WSEVLPOIQTETARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKALKIAGFNEP 512
DB 461 WSEVLPOIQTETARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKALKIAGFNEP 520
QY 513 NGNLOYGCKDITFEDFNFDQSTQNTKQALAEINATNYITVLDKILNANMILLIRDKRF 572
DB 521 NGNLOYGCKDITFEDFNFDQSTQNTKQALAEINATNYITVLDKILNANMILLIRDKRF 580
QY 573 HYDRNNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVI 632
DB 581 HYDRNNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVI 640

QY 633 NDRYDMLNISSLRQDKTFIDFKYNDKLPYISNPYKVNVAVTENTINPSENGDT 692
DB 641 NDRYDMLNISSLRQDKTFIDFKYNDKLPYISNPYKVNVAVTENTINPSENGDT 700
QY 693 STNGIKKILIFSKKGYEIG 711
DB 701 STNGIKKILIFSKKGYEIG 719
RESULT 8
PCT-US94-01624-31
Sequence 31, Application PC/TUS9401624
GENERAL INFORMATION:
APPLICANT: Leppia, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
TITLE OF INVENTION: RELATED METHODS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOMSEND and TOMSEND KHOURIE and CREW
STREET: Stewart Street Tower, 20th Floor, One Market
STREET: Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01624
FILING DATE: June 25, 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-31

Query Match 94.0%; Score 3430; DB 5; Length 719;
Best Local Similarity 91.9%; Pred. No. 9.8e-248;
Matches 679; Conservative 4; Mismatches 8; Indels 48; Gaps 3;

QY 1 EVKQENLLNESSSQGLLYGFFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60
DB 1 EVKQENLLNESSSQGLLYGFFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60
QY 61 QSAIWSGFIKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNS-----RKKSTSGPTVP 176
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNTATIMQKGNFLQSGTVP 180
QY 177 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSD 236
DB 181 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSD 240

QY 237 EKVTRIDKVNVPBARPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNSTTS 296
DB 241 EKVTRIDKVNVPBARPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNSTTS 300
QY 297 RTHT-----SNSNSTVAIDHSLAGERTWAETWGLNTAD 332
DB 301 RTHTSEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLAGERTWAETWGLNTAD 360
QY 333 TARLANIRYVNTGTAPIYVNLPTTSLVGLKQNTLAIKADENQLSQILAPNNYPSKNL 392
DB 361 TARLANIRYVNTGTAPIYVNLPTTSLVGLKQNTLAIKADENQLSQILAPNNYPSKNL 420
QY 393 APIALNAQKASSPTITANNYQFLEKTKQLRLDTQVYCNATYNFENGVRVDTGNS 452
DB 421 APIALNAQKASSPTITANNYQFLEKTKQLRLDTQVYCNATYNFENGVRVDTGNS 460
QY 453 WSEVLPQIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKAFGNENP 512
DB 461 WSEVLPQIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKAFGNENP 520
QY 513 NGNLOYQKGDITFEDFNDQDOTSONIKNQLAELNATNYTVLDKIKLNKQNTLIRDKRF 572
DB 521 NGNLOYQKGDITFEDFNDQDOTSONIKNQLAELNATNYTVLDKIKLNKQNTLIRDKRF 580
QY 573 HYDRNNIAVGADESIVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVI 632
DB 581 HYDRNNIAVGADESIVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVI 640
QY 633 NDRYDMLNIISSRQDKTFIDFKYNDKPLIYISNPNYKNNVAVTKENIINPSENGDT 692
DB 641 NDRYDMLNIISSRQDKTFIDFKYNDKPLIYISNPNYKNNVAVTKENIINPSENGDT 700
QY 693 STNGIKKILIFSKGYEIG 711
DB 701 STNGIKKILIFSKGYEIG 719

RESULT 9
US-09-273-839A-8
; Sequence 8, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick W
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 8
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-273-839A-8

Query Match 35.0%; Score 1275; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 2e-87;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 462 ETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKAFGNENPQYQK 521
DB 29 ETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKAFGNENPQYQK 88
QY 522 DITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLNKQNTLIRDKRFHYDRNNIAV 581
DB 89 DITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLNKQNTLIRDKRFHYDRNNIAV 148
QY 582 GADESIVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNI 641
DB 149 GADESIVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNI 208

QY 642 SSLROGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENIINPSENGDTSTNGIKKIL 701
DB 209 SSLROGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENIINPSENGDTSTNGIKKIL 268
QY 702 IFSKKGYEIG 711
DB 269 IFSKKGYEIG 278
RESULT 10
US-08-960-780-32
; Sequence 32, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Gerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narve, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schweits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA-708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 177C8
US-08-960-780-32

Query Match 20.8%; Score 759; DB 3; Length 881;
Best Local Similarity 29.7%; Pred. No. 3.5e-48;
Matches 236; Conservative 136; Mismatches 294; Indels 128; Gaps 32;
QY 4 QENRLNSESQGLGYFFSDLNQAPMVTSTGDLSPSSLEN--IFSENQYFQ 61
DB 39 QKQK---QKMDRKGILGYFFKQKDF-SNLTMEFAPTRDSTLIYDQQTANKLLDKQOEYQ 94
QY 62 SAIWSGFIKVKSDYEYTFATSNADNHYTWVDDQEVINKASNNKIRLEKGLYQIKIQY 121

Db 95 SIRWIGLQSKTGDFTNLSDEQAIIEINGKIIISNKGKQVHVHLEKGLVPIKIEYQ 154
QY 122 RENPTEKGLD-----FKLWYTDSONKKEVSSDNLQPLKOKXSS-----N 162
Db 155 SD--TKFNIDSKTFKELKFKIDSONQVQVQODELNFENFKESQEFLLAKPSKINLFT 212
QY 163 SRKSTSGAGTVPRDNDGIPDSLEVEGYTVVKNKRTFLSPWISNHEKGLTKYKSS 222
Db 213 QKMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264
QY 223 PEKWTASDPKSDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNT 282
Db 265 PLESHTVGDPTDYDEKAARDLDSNAKETFNPLVAAPPSVNVSMKVLSPNENLS---- 320
QY 283 DSETRISKNTSRTHTSNSS-----TVAIDHSLSLAGERTWAETWG- 327
Db 321 ----NSVESHSSTNWSYNTTEGASVEAGIGPKGIFSGVSVNYQHSETVAQE--WGTSTGN 374
QY 328 ---LNTADTARLANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPN 384
Db 375 TSQFNTASAGYLVANVYVNGTGAIYDVKPTTSFVL--NNDTIATITAKSNSTALNISP 433
QY 385 NYPSKNLAPALNAQKQDASSTPITMNYNQFLEKTKQLRLDQVTVGNATYNFENGR 444
Db 434 ESYPKKGQNGIAITSMDDFNSEHPTLNKKQVDNLLNNKPMMLTETQTDG---VYKIKDTH 490
QY 445 VRYDVTGNSWSEVLPOIQTARIIFNGKDLNLVERRIAAVNPSPDLETKPDMTLKALK 504
Db 491 GNIWVGWNGVLIQIKAKTASIIIVDDGB-RVAEKVAKDYENPEDKT-PSLTLKDALK 548
QY 505 IAP--GENEPNGNLQVQKDIETDFP--NFDQOTSONIKNQLAEL-----NATNIYVLD 555
Db 549 LSVYDEKEIEGLLYANKEPIYSSVMTYLDENTAKVTKQLNDTGTGKFDVSHLYDV-- 606
QY 556 KIKLNAKONILIRDKRPHYDRNNIAVGADESUVYKZAHREVINSSTEG-----L 603
Db 607 --KLTPRQNVTK--LSILYDN---AESDNGSIGKWTNTNIVSGNGNGKQYSSNNPDANL 660
QY 604 LLAND-----KDIRKILSGYVIEDTE-----GLKEVINRYDMLNT-- 641
Db 661 TLNDAQKLNKRDYVYISLMKSEKNTQCEITIDGIEIPIPTTKTVNKNKDYKRLDIIA 720
QY 642 SSLRQDKGTIDPKYNDKLPYISNPNYKVNVAVTKENTIIINPSENGDT--STNGIK-- 698
Db 721 HNIKSNIPISSIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSEIKQIYSRYGKLE 776
QY 699 -KLIFSKKGYEIG 711
Db 777 DGLIDKKGGINHYG 790

RESULT 11
US-09-073-898-32
; Sequence 32, Application US/C9073898
; Patent No. 6243669
; GENERAL INFORMATION:
; APPLICANT: Feteleson, Jerald S.
; APPLICANT: Schneft, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacey
; TITLE OF INVENTION: No. 6242669el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,898
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: NA-708C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 881 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: PSL177C8
US-09-073-898-32

Query Match 20.8%; Score 759; DB 3; Length 881;
Best Local Similarity 29.7%; Pred. No. 3.5e-48;
Matches 236; Conservative 136; Mismatches 294; Indels 128; Gaps 32;

QY 4 QENRLNSESSESSQGLGYFSDLNFOAPMVVTSITGDSIPSELEN--IPSENQYFQ 61
Db 39 QKNQ--QKEMDRKGLGYFKGDF--SNLTFAPTRDSTLIYDQOTANKLLDKKQEQY 94
QY 62 SAIWGFKVKKSDBYTATSAADNHVTWVDDQEVINKASNSKIRLEKGLVQIKIQY 121
Db 95 SIRWIGLQSKTGDFTNLSDEQAIIEINGKIIISNKGKQVHVHLEKGLVPIKIEYQ 154
QY 122 RENPTEKGLD-----FKLWYTDSONKKEVSSDNLQPLKOKXSS-----N 162
Db 155 SD--TKFNIDSKTFKELKFKIDSONQVQVQODELNFENFKESQEFLLAKPSKINLFT 212
QY 163 SRKSTSGAGTVPRDNDGIPDSLEVEGYTVVKNKRTFLSPWISNHEKGLTKYKSS 222
Db 213 QKMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264
QY 223 PEKWTASDPKSDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNT 282
Db 265 PLESHTVGDPTDYDEKAARDLDSNAKETFNPLVAAPPSVNVSMKVLSPNENLS---- 320
QY 283 DSETRISKNTSRTHTSNSS-----TVAIDHSLSLAGERTWAETWG- 327
Db 321 ----NSVESHSSTNWSYNTTEGASVEAGIGPKGIFSGVSVNYQHSETVAQE--WGTSTGN 374
QY 328 ---LNTADTARLANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPN 384
Db 375 TSQFNTASAGYLVANVYVNGTGAIYDVKPTTSFVL--NNDTIATITAKSNSTALNISP 433
QY 385 NYPSKNLAPALNAQKQDASSTPITMNYNQFLEKTKQLRLDQVTVGNATYNFENGR 444
Db 434 ESYPKKGQNGIAITSMDDFNSEHPTLNKKQVDNLLNNKPMMLTETQTDG---VYKIKDTH 490

QY 445 VRDTGSNWSEVLPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALK 504
DB 491 GNIVTGGWNGVIOQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPEKDT-PSLTLDKALK 549
QY 505 IAF--GFNEPNNGLOQKIDTEPFD--NFDQOTSQNIKNQLAEL-----NATNIYTVLD 555
DB 549 LSPYDEIKEIEGLLYKKNPIYESSVMTYLDNTAKETVKQLNDTTGKFDVSHLYDV-- 606
QY 556 KIKLNAKMNIILDRKFRHYDRNNIAVGADESVMKAEHREVINSSTEG-----L 603
DB 607 --KLTPKNNVTIK-LSILYDN--AESNDNSIGKWTNTNIVSGGNGKKQYSSNNPDANL 660
QY 604 LLNID-----KDIRKILSGYVIEDETE-----GLKEVINDRYDMLNI-- 641
DB 661 TLNTDAQEKLNKRDYIISLYMKSEKNTQCEITIDGEIYPIITTKVNVNKNYKELDIIA 720
QY 642 SSLRQDGKTFIDFKYNDKLPYISNPNKYKNVAVTKENTINPSENGDT-STNGIK-- 698
DB 721 HNKENPSSIIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSEIKQIYSRYGKLE 776
QY 699 -KILIFSKGYEIG 711
DB 777 DGLIDKKGGIHYG 790

RESULT 12
US-09-307-106-8
Sequence 8, Application US/09307106
Patent No. 6603063
GENERAL INFORMATION:
APPLICANT: Fetteleson, Jerald S.
APPLICANT: Schnepf, H. Ernest
APPLICANT: Narva, Kenneth E.
APPLICANT: Stockhoff, Brian A.
APPLICANT: Schmeits, James
APPLICANT: Loewer, David
APPLICANT: Dullum, Charles Joseph
APPLICANT: Muller-Cohn, Judy
APPLICANT: Stamp, Lisa
APPLICANT: Morrill, George
APPLICANT: Finstad-Lee, Stacey
TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide
TITLE OF INVENTION: Sequences Which Encode These Toxins
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/307,106
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/073,898
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 981 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE: PS177C8a
INDIVIDUAL ISOLATE: PS177C8a
US-09-307-106-8

Query Match 20.8%; Score 759; DB 4; Length 881;
Best Local Similarity 29.7%; Pred. No. 3.5e-48;
Matches 236; Conservative 136; Mismatches 294; Indels 128; Gaps 32;

QY 4 QENRLNSESSESSQGLLYGYSFSLNTOAPMVVTSSTGDLSPSSLEN--IFSENQYFQ 61
DB 39 QKQK--QKEMDRKGLLYGFKGDF-SNLTMPFAPTRDSTLIYDQOTANKLLDKKQYEQ 94
QY 62 SAIWSGPIKVKSDDEVTEATSADNHVTMWVDDOEIVNKASNSNKIRLEKGRLYQIKIYQ 121
DB 95 SIRWIGLIQSKETGDFTFNLSEDEQALIBINGKLIISNKGKEQVHLEKGLVPKIEYQ 154
QY 122 RENPTEKGLD-----FKLYWDSQNKKEVISSDNLQLPKOKSS-----N 162
DB 155 SD--TKFNIDSKTFKELKLFKIDSONQVQVQOQDELNPFNFKEKSEQEFLLAKPSKINLFT 212
QY 163 SRKRSTSGAPTPDRDNDGIPDSLEVEGYTVGVNKRITFLSPWISNIHEKGLTKYKSS 222
DB 213 QKWKREIDED---TDTGDSIPDLWEENGTYI---QNRILAVKWDLSL-ASKGYTFVSN 264
QY 223 PEKWTASDPYSPPEKVTGRIDKNVSPARHPHVAAYPIVHVDMENIILSKNDQSTQNT 282
DB 265 PLESHTVGDPTDYKAAARDLDSNAKETENPLVAAFPSPVNVSMKVVILSPENLS--- 320
QY 283 DSETETISKNTSRTHTSNSNS-----TVAIDHLSLSLAGERTWAETMG- 327
DB 321 ----NSVESSTNWSVTNTEGASVEAGIGPKGISFGVSVNYQHSEVAGE--WGTSTGN 374
QY 328 ---LNTADTARLANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQAIIAPN 384
DB 375 TSQENTASAGYLNANRVYNNVGTGAIYDKVPTTSFVL--NNDTIATITAKSNSTALNISPG 433
QY 385 NYFSPKMLAPIALNAQKQDASPTITMNYNQFLELEKTKQLRLDPTDQVYGNIAVYFNGR 444
DB 434 ESYFKGQNGIAITSMDDFNSHPITLNRKQVDNLLNNKPMLENTQDTG---VYKIDTH 490
QY 445 VRDTGSNWSEVLPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALK 504
DB 491 GNIVTGGWNGVIOQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPEKDT-PSLTLDKALK 548
QY 505 IAF--GFNEPNNGLOQKIDTEPFD--NFDQOTSQNIKNQLAEL-----NATNIYTVLD 555
DB 549 LSPYDEIKEIEGLLYKKNPIYESSVMTYLDNTAKETVKQLNDTTGKFDVSHLYDV-- 606
QY 556 KIKLNAKMNIILDRKFRHYDRNNIAVGADESVMKAEHREVINSSTEG-----L 603
DB 607 --KLTPKNNVTIK-LSILYDN--AESNDNSIGKWTNTNIVSGGNGKKQYSSNNPDANL 660
QY 604 LLNID-----KDIRKILSGYVIEDETE-----GLKEVINDRYDMLNI-- 641
DB 661 TLNTDAQEKLNKRDYIISLYMKSEKNTQCEITIDGEIYPIITTKVNVNKNYKELDIIA 720
QY 642 SSLRQDGKTFIDFKYNDKLPYISNPNKYKNVAVTKENTINPSENGDT-STNGIK-- 698
DB 721 HNKENPSSIIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSEIKQIYSRYGKLE 776
QY 699 -KILIFSKGYEIG 711

Db 777 DGLIDKKGIIHYG 790

RESULT 13
US-09-850-351A-32
Sequence 32, Application US/09850351A
Patent No. 6656908
GENERAL INFORMATION:
APPLICANT: Feltelson, Gerald S.
Schnepf, H. Ernest
Narva, Kenneth B.
Stockhoff, Brian A.
Schmeits, James
Loewer, David
Dullum, Charles Joseph
Muller-Cohn, Judy
Stamp, Lisa
Morrill, George
TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide
Sequences Which Encode These Toxins
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/850,351A
FILING DATE: 07-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/073,898
FILING DATE: 06-MAY-1998
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: WA-708CD1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 881 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: PS177C8
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-850-351A-32
Query Match 20.8%; Score 759; DB 4; Length 881;
Best Local Similarity 29.7%; Pred. No. 3.5e-48;
Matches 236; Conservative 136; Mismatches 294; Indels 128; Gaps 32;
Qy 4 QENLLNESSSQGLCYFSDLNFPQPMVVTSTGDISIPSSLEN--IPSENYQYQ 61
Db 39 QKNQ---QKENDRKGLLGYFKGKDF-SNLTMFAPTRDSTLIYDQQTANKLLDKKQOEYQ 94
Qy 62 SAIVSGFLVKVKSDEYTFATSDADNHVTWVDQEVINKASNSKIRLEKRLVQIKYQY 121

95 SIRWIGLIQSKETGDTFNLSEDEQAIIIEINGKIIISNKGKEQVHVHLEKGLVPIKIEYQ 154
122 RENPTEKGLD-----FKLYWTDSONKKEVIGSDNLQLPELKQKSS-----N 162
155 SD--TKFNIDSKTFKELKLFKIDSQNOQVQOQDELBNPEFNKESQBEFLAKPSKINLPT 212
163 SRKRSSTAGPTVDRDNDGIPDSLEVBGYTVDVVKNRKRTFLSPWISNHEKGLTKYKSS 222
213 QKMKREIDED--TDTGDSIPDLMEENGYYTI---QNRIAVKWDDSL-ASKGYTKFVSN 264
223 PERKSTASDPYSDPEKVTGRIDKQVSPPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNT 282
265 PLESHTVGDPYTDYERKAARDLDSNAKETFNPLVAAPSVYVSMKVVILSPENLS---- 320
283 DSETRTISKNSTSRTHTSNNS-----TVAIDHSLSLAGERTWAETMG- 327
321 ---NSVESHSSTNWSYNTTEGASVEAGIGPKGISFGVSNYQHSSTVAQE--WGTSTGN 374
328 ---LNTADTARLANIRVYNTGTAPYNYVLPTTSLVGLKNOTLATIKADENOLSQLAPN 384
375 TSQFNTASAGYLNANRVNNVGTGALYDVKPTTFSVL--NDIITATITAKSNSTALNISP 433
385 NYTPSKNLAPIALNAQKASSTPTIMYNYQFLEKTKQLRLDQDQVYGNIAIYNFENGR 444
434 ESYFKKGQNGIAITSMDDPNSHEITLNXKQVDNLLANKPMLENTQTDG---VYKIDTH 490
445 VRVDTGSNWSEVLPOIQTETARIIFNGKOLNIVERIAAVNPSPDPLETTKPDMLKEALK 504
491 GNVITGGEWNGVIOQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPEDKT-PSLTLDKALK 548
505 IAF--GFNEPENGVLQYQCKDITEPDF--NFDQOTSQNIKNQLAEL-----NATNIYTVLD 555
549 LSYPDDEIKETGLLYYKKNKPIYESSVMTYLDENTAKEVTKQLNDTTGKFKDVSHLYDV-- 606
556 KIKUNAKONILIRDRFHYDRNNIAGADESVVKEAHRREVINSSSTEG-----L 603
607 --KLTPKQNVTK-LSLYDN--AESNDNSICKWTNTIVSGNKKQYSSNNDANL 660
604 LLNID-----KDIRKILSGVIVEIDTE-----GLKEVINDRYDMLNI-- 641
661 TLNTDAQEKLNKRDYIISLYMKSEKNTQCEIIDGIEIYITTKTVNVNKNKDYKRLDIIA 720
642 SSLRQDQKTFIDFKYKNDKPLIYISNPNYKNVYAVTKENTINPSENGDT-STNGIK-- 698
721 HNIKSNPISSIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSSBIKQIYSRYGIKLE 776

Db 699 -KILIFSKGYEIG 711
777 DGLIDKKGIIHYG 790

RESULT 14
US-08-471-033-5
Sequence 5, Application US/08471033
Patent No. 5770696
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA

! INFORMATION FOR SEQ ID NO: 5:

! SEQUENCE CHARACTERISTICS:

! LENGTH: 884 amino acids

! TYPE: amino acid

! TOPOLOGY: linear

! MOLECULE TYPE: protein

US-08-471-044-5

```
Query Match      20.8%; Score 757; DB 2; Length 884;
Best Local Similarity 29.7%; Pred. No. 5e-48;
Matches 238; Conservative 131; Mismatches 289; Indels 144; Gaps 32;

QY 4 QENRLINESSESSQGLGYFSDLNFQAPMVVTSSTTGDLSIPSSSELEN--IPSENQYFQ 61
Db 42 QKNQ--QKEMDRKGLGYFPKGDYF--SNLTWFAPTRDSTLIYDQOTANKLLDKKQBYQ 97
QY 62 SAIWGFIKVKSDSYTATSDNHTVMVDQEVINKASNSNKIRLEKGRLYQIKIQYQ 121
Db 98 SIWIGLIQSKETGDTFNLSDEQAIIIEINGKIISNKGKEKQVVHLEKGLVPIKIEYQ 157
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISNDNLQLPKQKSS-----N 162
Db 158 SD--TKFNIDSNTFKELKLFKIDSONPOQVQODELRNPFNPKESQEFLLAKPSKINLFT 215
QY 163 SRKKSTSGAPVPRDNDGIPDSLEVEGYTVDVKNKTFFLSPWISNIHEKKGLTKYKSS 222
Db 216 QMKKEIDEDE--TDTGDSIPDLWEENGYYTI---QNRIVAKWDDSL-ASKGYTKFVSN 267
QY 223 PEKMWSTADPYGDFEKVTGRIDKNVSPEARHPLVAAPYIVHVDVMEIILSKNEDOSTQNT 282
Db 268 PLESHTVGDPYTDYEKAARDLDSNAKETFNPLVAAPFVSVNVMKSVILSPENLS---- 323
QY 283 DSETRTISKNTSRTSTNSNSS-----TVALDHSLSLAGERTWAETWG- 327
Db 324 ---NSVESHSSTNWSYTTNTGASVEAGIGPKGISFGVSVNYQHSSETVAQE--WGTSTGN 377
QY 328 ---LNTADTARLNANRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPN 384
Db 378 TSQFNTASAGYLNANRYVNVGTGALYDVKPTTSFVL--NNDTIATITAKSNSTALN1SPG 436
QY 385 NYYPKKNLAPIANAQKADASSPTTNVYNQFLEKTKQLRLDQDVYGNATYNFENGR 444
Db 437 ESYFKEKGQNGIAITSDDFNSHPITLNNKQVDNLLNNKPMMLTNTQTDG---VYKIKDTH 493
QY 445 VRYDTGSNKEVLPOIQETTARIIFNGKDLNLVERRIAAVNPSPDPLETTKPDMTLKEALK 504
Db 494 GNVITGGEWNGVTQIQAKTASLIYDDGE--RVAEKRAAKOYENPEDKT--PSLTUKDALK 551
QY 505 IAF--GFNEPNGNLOYQKGDITEFDF--NFDQOTSQNIKNQLAEL-----NATNIYTVLD 555
Db 552 LSYPDKEIKETEGLLYKKNKPIYESSVMTYLDENTAKEVTQLNDTGTGKFDVSHLYDV-- 609
QY 556 KIKLNAMNILLRDKPFDNRNNIAVGADESUVVKEAHREVINSSTEG-----L 603
Db 610 --KLTPKMNVTIK--LSILYDN---AESNDNSIGKWTNTNIVSGGNNKKQYSSNNPDANL 663
QY 604 LLNID-----KDIRKILSGYIIVEIDTE-----GLKEVINDRYDMLN--- 640
Db 664 TLNTDAQEKLNKRDYVYISLYMKSEKNTQCEITDGEIYPIITTKTVNVKNKYKELDIIA 723
QY 641 -----ISSLRQCGKTFDFKKYNDKPLPYISNPYKNVYAVTKENTIINPSENGDT- 692
Db 724 HNTKSNPISSLH-----IKTNDBITLFWDDISI--TDVASIKPEN--LTDSEIKQIY 771
QY 693 STNGIK---KILIFSKKGYEIG 711
Db 772 SRYGKLEGDGILLDKGGIHVG 793
```

Search completed: May 3, 2004, 19:42:50
Job time : 17.6056 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:40:13 ; Search time 35.4429 Seconds
(without alignments)
5560.545 Million cell updates/sec

Title: US-09-848-909A-16
Perfect score: 3648
Sequence: 1 EVKQENRLNSESSESSQGLL.....TSTNGIKILIFSKGYEIG 711

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/US07_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3607	98.9	735	15	US-10-410-647-30
2	3607	98.9	736	12	US-09-848-909-1
3	3607	98.9	736	12	US-09-848-909-2
4	3607	98.9	736	12	US-09-848-909-3
5	3607	98.9	736	12	US-09-848-909-4
6	3607	98.9	736	12	US-09-848-909-5
7	3607	98.9	736	12	US-09-848-909-6
8	3607	98.9	736	12	US-09-848-909-7
9	3607	98.9	736	12	US-09-848-909-8
10	3607	98.9	736	12	US-09-848-909-9
11	3607	98.9	736	12	US-09-848-909-10
12	3607	98.9	736	12	US-09-848-909-11
13	3607	98.9	736	12	US-09-848-909-12
14	3607	98.9	736	12	US-09-848-909-13
15	3607	98.9	736	12	US-09-848-909-14

16	3607	98.9	736	12	US-09-848-909-15	Sequence 15, Appl
17	3607	98.9	736	12	US-09-848-909-16	Sequence 16, Appl
18	3607	98.9	736	12	US-09-848-909-17	Sequence 17, Appl
19	3607	98.9	736	12	US-09-848-909-18	Sequence 18, Appl
20	3607	98.9	736	12	US-09-848-909-19	Sequence 19, Appl
21	3607	98.9	736	12	US-09-848-909-20	Sequence 20, Appl
22	3607	98.9	736	12	US-09-848-909-21	Sequence 21, Appl
23	3607	98.9	736	12	US-09-848-909-22	Sequence 22, Appl
24	3607	98.9	736	12	US-10-442-502-7	Sequence 7, Appl
25	3607	98.9	736	15	US-10-442-502-5	Sequence 5, Appl
26	3607	98.9	764	15	US-10-442-502-6	Sequence 6, Appl
27	3604	98.8	735	12	US-10-402-466A-9	Sequence 9, Appl
28	3604	98.8	735	12	US-10-402-466A-13	Sequence 13, Appl
29	3604	98.8	735	14	US-10-332-282-13	Sequence 13, Appl
30	3600	98.7	735	12	US-09-848-909-30	Sequence 30, Appl
31	3600	98.7	764	12	US-10-253-286-681	Sequence 681, App
32	3600	98.7	764	15	US-10-245-871-681	Sequence 681, App
33	3586	98.3	764	9	US-09-747-521-4	Sequence 4, Appl
34	3586	98.3	764	13	US-10-106-014-4	Sequence 4, Appl
35	3586	98.3	764	13	US-10-105-695-4	Sequence 4, Appl
36	3586	98.3	764	14	US-10-105-694-4	Sequence 4, Appl
37	2891	79.4	599	12	US-09-848-909-24	Sequence 24, Appl
38	2891	79.2	595	14	US-10-332-282-11	Sequence 11, Appl
39	2764	75.8	573	12	US-10-402-466A-22	Sequence 22, Appl
40	2752	75.4	569	15	US-10-442-502-8	Sequence 8, Appl
41	2342	64.2	487	14	US-10-332-282-5	Sequence 5, Appl
42	2077	56.9	423	12	US-10-402-466A-24	Sequence 24, Appl
43	2027	55.6	426	14	US-10-332-282-9	Sequence 9, Appl
44	1478	40.5	318	14	US-10-332-282-7	Sequence 7, Appl
45	1338	36.7	258	14	US-10-332-282-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-410-647-30
; Sequence 30, Application US/10410647
; Publication No. US20030235818A1
; GENERAL INFORMATION:
; APPLICANT: PLEXUS VACCINE, INC.
; APPLICANT: Katritch, Vaevoled
; APPLICANT: Bordiner, Andrew
; APPLICANT: Deans, Robert
; APPLICANT: Summer, Mary
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES, AND METHOD OF IDENTIFYING SAME
; FILE REFERENCE: PLEX1110-1
; CURRENT APPLICATION NUMBER: US/10/410,647
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/373,668
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/371,256
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/371,250
; PRIOR FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 30
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-410-647-30

Query Match 98.9%; Score 3607; DB 15; Length 735;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;
QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSITGDLSPSELENIPSENYQYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSITGDLSPSELENIPSENYQYF 60
QY 61 QSAWSGFIKYKGDYTFATSDAHVMTWYDDQEVINKASNSKIRLEKRLVQIKIQY 120

Db 61 QSAIWSGFIKVKSDSEYTFATSADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
 QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELKOKSSNRKRSSTAGPTVPDRDN 180
 Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELKOKSSNRKRSSTAGPTVPDRDN 180
 QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDPYSDFEKVT 240
 Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDPYSDFEKVT 240
 QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
 Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
 QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
 Db 301 SEVHGNAEVHASFPIIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
 QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA 396
 Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA 420
 QY 397 LNAQKADASTPTIMYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA 456
 Db 421 LNAQKADASTPTIMYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA 480
 QY 457 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 516
 Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 540
 QY 517 QYQKDIITEFDNFDOQTSONIKQNLAEINATNITVLDKIKLNAKNNILIRDKRPHYDR 576
 Db 541 QYQKDIITEFDNFDOQTSONIKQNLAEINATNITVLDKIKLNAKNNILIRDKRPHYDR 600
 QY 577 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 636
 Db 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 QY 637 DMLNSSLRODQKTFIDFKYNDKPLXIISNPNKYVNVYAVTKENTIIINPSENGDSTNG 696
 Db 661 DMLNSSLRODQKTFIDFKYNDKPLXIISNPNKYVNVYAVTKENTIIINPSENGDSTNG 720
 QY 697 IKKILIFSCKGYEIG 711
 Db 721 IKKILIFSCKGYEIG 735

RESULT 2

US-09-848-909-1
 ; Sequence 1, Application US/09848909
 ; Publication No. US20020039588A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Collier, R. John
 ; APPLICANT: Sellman, Brett R.
 ; TITLE OF INVENTION: Compounds and Methods for the Treatment
 ; TITLE OF INVENTION: and Prevention of Bacterial Infection
 ; FILE REFERENCE: 00742/060002
 ; CURRENT APPLICATION NUMBER: US/09/848,909
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: US 60/201,800
 ; PRIOR FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 736
 ; TYPE: PRT
 ; ORGANISM: Bacillus anthracis
 US-09-848-909-1

Query Match 98.9%; Score 3607; DB 12; Length 736;
 Best Local Similarity 96.3%; Pred. No. 1.2e-262;
 Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLINESSSQGLGYFSDLNQFQPMVVTSTTGDLSIPSSSELENIIPSENOVF 60
 Db 1 EVKQENRLINESSSQGLGYFSDLNQFQPMVVTSTTGDLSIPSSSELENIIPSENOVF 60
 QY 61 QSAIWSGFIKVKSDSEYTFATSADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
 Db 61 QSAIWSGFIKVKSDSEYTFATSADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
 QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELKOKSSNRKRSSTAGPTVPDRDN 180
 Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELKOKSSNRKRSSTAGPTVPDRDN 180
 QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDPYSDFEKVT 240
 Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDPYSDFEKVT 240
 QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
 Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
 QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
 Db 301 SEVHGNAEVHASFPIIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
 QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA 396
 Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA 420
 QY 397 LNAQKADASTPTIMYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA 456
 Db 421 LNAQKADASTPTIMYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA 480
 QY 457 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 516
 Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 540
 QY 517 QYQKDIITEFDNFDOQTSONIKQNLAEINATNITVLDKIKLNAKNNILIRDKRPHYDR 576
 Db 541 QYQKDIITEFDNFDOQTSONIKQNLAEINATNITVLDKIKLNAKNNILIRDKRPHYDR 600
 QY 577 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 636
 Db 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 QY 637 DMLNSSLRODQKTFIDFKYNDKPLXIISNPNKYVNVYAVTKENTIIINPSENGDSTNG 696
 Db 661 DMLNSSLRODQKTFIDFKYNDKPLXIISNPNKYVNVYAVTKENTIIINPSENGDSTNG 720
 QY 697 IKKILIFSCKGYEIG 711
 Db 721 IKKILIFSCKGYEIG 735

RESULT 3

US-09-848-909-2
 ; Sequence 2, Application US/09848909
 ; Publication No. US20020039588A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Collier, R. John
 ; APPLICANT: Sellman, Brett R.
 ; TITLE OF INVENTION: Compounds and Methods for the Treatment
 ; TITLE OF INVENTION: and Prevention of Bacterial Infection
 ; FILE REFERENCE: 00742/060002
 ; CURRENT APPLICATION NUMBER: US/09/848,909
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: US 60/201,800
 ; PRIOR FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 736
 ; TYPE: PRT
 ; ORGANISM: Bacillus anthracis

US-09-848-909-2

Query Match
Best Local Similarity 98.9%; Score 3607; DB 12; Length 736;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNESSSQGLLYYFSDLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSQGLLYYFSDLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNQLPQLKQKSSNRKGRSTASGTPVDPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNQLPQLKQKSSNRKGRSTASGTPVDPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFSPWISNIHEKKGLTKYKSSPEKWSASDPSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFSPWISNIHEKKGLTKYKSSPEKWSASDPSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRISKNSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVSFFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVLTPTSLVLGKNOTLATIKADENQLSOLILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVLTPTSLVLGKNOTLATIKADENQLSOLILAPNNYPSKNLAPIA 420
Qy 397 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVFNENGRVVRVDTGNSWSEV 456
Db 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVFNENGRVVRVDTGNSWSEV 480
Qy 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETKPDMTLKALKIAFGFNEPKNL 516
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETKPDMTLKALKIAFGFNEPKNL 540
Qy 517 QYQCKDITEFDNFDDQTSQNIKNQALAEINATNIYVLDKIKNAKMILIRDKRFHYDR 576
Db 541 QYQCKDITEFDNFDDQTSQNIKNQALAEINATNIYVLDKIKNAKMILIRDKRFHYDR 600
Qy 577 NNIAVGADESUVKEAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 637 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIINPSENGTSTNG 696
Db 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIINPSENGTSTNG 720
Qy 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735

RESULT 4

US-09-848-909-3
; Sequence 3, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848, 909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-09-848-909-3

Query Match

Best Local Similarity 98.9%; Score 3607; DB 12; Length 736;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNESSSQGLLYYFSDLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSQGLLYYFSDLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNQLPQLKQKSSNRKGRSTASGTPVDPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNQLPQLKQKSSNRKGRSTASGTPVDPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFSPWISNIHEKKGLTKYKSSPEKWSASDPSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFSPWISNIHEKKGLTKYKSSPEKWSASDPSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRISKNSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVSFFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVLTPTSLVLGKNOTLATIKADENQLSOLILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVLTPTSLVLGKNOTLATIKADENQLSOLILAPNNYPSKNLAPIA 420
Qy 397 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVFNENGRVVRVDTGNSWSEV 456
Db 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVFNENGRVVRVDTGNSWSEV 480
Qy 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETKPDMTLKALKIAFGFNEPKNL 516
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETKPDMTLKALKIAFGFNEPKNL 540
Qy 517 QYQCKDITEFDNFDDQTSQNIKNQALAEINATNIYVLDKIKNAKMILIRDKRFHYDR 576
Db 541 QYQCKDITEFDNFDDQTSQNIKNQALAEINATNIYVLDKIKNAKMILIRDKRFHYDR 600
Qy 577 NNIAVGADESUVKEAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 637 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIINPSENGTSTNG 696
Db 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIINPSENGTSTNG 720
Qy 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735

RESULT 5

US-09-848-909-4
; Sequence 4, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 736
; TYPE: PRN
; ORGANISM: Bacillus anthracis
US-09-848-909-4

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Qy 61 QSAIWSGFTKVKKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFTKVKKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

Qy 121 QRENPEKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNSRKRSTSGAPVDPDRN 180
Db 121 QRENPEKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNSRKRSTSGAPVDPDRN 180

Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSWSTASDPYDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSWSTASDPYDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVAHSPFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 337 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

Qy 397 LNAQKDSASPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVVDTSNWSSEV 456
Db 421 LNAQDDFSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVVDTSNWSSEV 480

Qy 457 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 516
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540

Qy 517 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIVTLDKIKLNAKMILLIRDKRFHYDR 576
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIVTLDKIKLNAKMILLIRDKRFHYDR 600

Qy 577 NNIAVGADESUVKEAREHREINVESTGELLNIDKDKRILSGYIVIEDTEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAREHREINVESTGELLNIDKDKRILSGYIVIEDTEGLKEVINDRY 660

Qy 637 DMLNLSLRQDGKTFIDFKKYNKDKPLYSISNPNYKVNVAVKENTIIINPSNGDSTNG 696
Db 661 DMLNLSLRQDGKTFIDFKKYNKDKPLYSISNPNYKVNVAVKENTIIINPSNGDSTNG 720

Qy 697 IKKILFSSKGYEIG 711
Db 721 IKKILFSSKGYEIG 735

RESULT 6
US-09-848-909-5

; Sequence 5, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Cellier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRN
; ORGANISM: Bacillus anthracis
US-09-848-909-5

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Qy 61 QSAIWSGFTKVKKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFTKVKKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

Qy 121 QRENPEKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNSRKRSTSGAPVDPDRN 180
Db 121 QRENPEKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNSRKRSTSGAPVDPDRN 180

Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSWSTASDPYDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSWSTASDPYDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVAHSPFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 337 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

Qy 397 LNAQKDSASPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVVDTSNWSSEV 456
Db 421 LNAQDDFSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVVDTSNWSSEV 480

Qy 457 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 516
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540

Qy 517 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIVTLDKIKLNAKMILLIRDKRFHYDR 576
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIVTLDKIKLNAKMILLIRDKRFHYDR 600

Qy 577 NNIAVGADESUVKEAREHREINVESTGELLNIDKDKRILSGYIVIEDTEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAREHREINVESTGELLNIDKDKRILSGYIVIEDTEGLKEVINDRY 660

Qy 637 DMLNLSLRQDGKTFIDFKKYNKDKPLYSISNPNYKVNVAVKENTIIINPSNGDSTNG 696
Db 661 DMLNLSLRQDGKTFIDFKKYNKDKPLYSISNPNYKVNVAVKENTIIINPSNGDSTNG 720

Qy 697 IKKILFSSKGYEIG 711

Db 721 IKKILFSKGYEIG 735

RESULT 7

US-09-848-909-6

; Sequence 6, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-6

Query Match 98.9%; Score 3607; DB 12; Length 736;

Best Local Similarity 96.3%; Pred. No. 1.2e-262;

Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Qy 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKY 120

Db 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRSTASDPSDFEKT 180

Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRSTASDPSDFEKT 180

Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPSDFEKT 240

Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPSDFEKT 240

Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336

Db 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 337 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396

Db 361 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

Qy 397 LNAQKASSTPIITMNNQFLEBKTKQLRLDTPQVYGNITATYNFENGRVRVDTGSNWSEV 456

Db 421 LNAQDDFSSPTITMNNQFLEBKTKQLRLDTPQVYGNITATYNFENGRVRVDTGSNWSEV 480

Qy 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNPNGL 516

Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNPNGL 540

Qy 517 QYQKDIETEFDFNFDOQTSONIKNQALAEALNATNIYTVLDKIKLNAKMNTLIRDKRPHYDR 576

Db 541 QYQKDIETEFDFNFDOQTSONIKNQALAEALNATNIYTVLDKIKLNAKMNTLIRDKRPHYDR 600

Qy 577 NNIAVGADESIVKAEHREVNSTEGLLNIDKIDKILSGYVEIETEGEKEVINDRY 636

Db 601 NNIAVGADESIVKAEHREVNSTEGLLNIDKIDKILSGYVEIETEGEKEVINDRY 660

Qy 637 DMLNSSLRODQKTFIDFKKYNDKLPYISNPYKNTYAVTKENTIIINPSENGDTSTNG 696

Db 661 DMLNSSLRODQKTFIDFKKYNDKLPYISNPYKNTYAVTKENTIIINPSENGDTSTNG 720

Qy 697 IKKILFSKGYEIG 711

Db 721 IKKILFSKGYEIG 735

RESULT 8

US-09-848-909-7

; Sequence 7, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-7

Query Match 98.9%; Score 3607; DB 12; Length 736;

Best Local Similarity 96.3%; Pred. No. 1.2e-262;

Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Qy 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKY 120

Db 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRSTASDPSDFEKT 180

Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRSTASDPSDFEKT 180

Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPSDFEKT 240

Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPSDFEKT 240

Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336

Db 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 337 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396

Db 361 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

Qy 397 LNAQKASSTPIITMNNQFLEBKTKQLRLDTPQVYGNITATYNFENGRVRVDTGSNWSEV 456

Db 421 LNAQDDFSSPTITMNNQFLEBKTKQLRLDTPQVYGNITATYNFENGRVRVDTGSNWSEV 480

Qy 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNPNGL 516

Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNPNGL 540

Qy 517 QYQKDIETEFDFNFDOQTSONIKNQALAEALNATNIYTVLDKIKLNAKMNTLIRDKRPHYDR 576

```
Db 541 QYQKDIETFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRFHYDR 600
QY 577 NNIAGADESVVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEBTBGLKEVINDRY 636
Db 601 NNIAGADESVVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEBTBGLKEVINDRY 660
QY 637 DMLNIISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVVAVTKENTIIINPSENGDTSTNG 696
Db 661 DMLNIISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735

RESULT 9
US-09-848-909-8
; Sequence 8, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-8

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGLDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGLDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQSSNSRKRKSTASDPSDFEKT 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQSSNSRKRKSTASDPSDFEKT 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKSTASDPSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKSTASDPSDFEKT 240
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAETMGLNTADTARL 336
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAETMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYVLPFTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVLPFTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 397 LNAQKQDASSTPIITWYVNFQLEKTKQLRLDITDQVYGNATYNFENGRVVRVDTGSNWSEV 456
Db 421 LNAQKQDSSSTPIITWYVNFQLEKTKQLRLDITDQVYGNATYNFENGRVVRVDTGSNWSEV 480
```

```
QY 457 LPOIQETTARIIPNGKDLNLVERRIAAVNPSPDLETTKPDWTLKEALKIATFGNEPENG 516
Db 481 LPOIQETTARIIPNGKDLNLVERRIAAVNPSPDLETTKPDWTLKEALKIATFGNEPENG 540
QY 517 QYQKDIETFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRFHYDR 576
Db 541 QYQKDIETFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRFHYDR 600
QY 577 NNIAGADESVVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEBTBGLKEVINDRY 636
Db 601 NNIAGADESVVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEBTBGLKEVINDRY 660
QY 637 DMLNIISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVVAVTKENTIIINPSENGDTSTNG 696
Db 661 DMLNIISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735

RESULT 10
US-09-848-909-9
; Sequence 9, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-9

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGLDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGLDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQSSNSRKRKSTASDPSDFEKT 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQSSNSRKRKSTASDPSDFEKT 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKSTASDPSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKSTASDPSDFEKT 240
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAETMGLNTADTARL 336
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAETMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYVLPFTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 337 NANIRYVNTGTAPIYVLPFTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
```



```
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420
Qy 397 LNAQKADASTPTTMYNQFLELEKTKQLRLDQVYGNIAATYFNGRVRVDTGNSWSEV 456
Db 421 LNAQDDFSTPTTMYNQFLELEKTKQLRLDQVYGNIAATYFNGRVRVDTGNSWSEV 480
Qy 457 LPOIQTETARIIFNGKDLNLRERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 516
Db 481 LPOIQTETARIIFNGKDLNLRERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
Qy 517 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 576
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 600
Qy 577 NNIAGADESVVKEAAREVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 636
Db 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 637 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNYKVNYAVTKENTIINPSENGDTSTNG 696
Db 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNYKVNYAVTKENTIINPSENGDTSTNG 720
Qy 697 IKKILIFSCKGYEIG 711
Db 721 IKKILIFSCKGYEIG 735

RESULT 11
US-09-848-909-10
; Sequence 10, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-10

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPELKOKGSNNRKKRSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPELKOKGSNNRKKRSTAGPTVPDRDN 180
Qy 181 DGIPLSLEVEGTVVVKNNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYDFEKVT 240
Db 181 DGIPLSLEVEGTVVVKNNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYDFEKVT 240
Qy 241 GRIDKXVSPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKXVSPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTSRHT 300
```

```
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHAFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILLAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420
Qy 397 LNAQKADASTPTTMYNQFLELEKTKQLRLDQVYGNIAATYFNGRVRVDTGNSWSEV 456
Db 421 LNAQDDFSTPTTMYNQFLELEKTKQLRLDQVYGNIAATYFNGRVRVDTGNSWSEV 480
Qy 457 LPOIQTETARIIFNGKDLNLRERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 516
Db 481 LPOIQTETARIIFNGKDLNLRERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
Qy 517 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 576
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 600
Qy 577 NNIAGADESVVKEAAREVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 636
Db 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 637 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNYKVNYAVTKENTIINPSENGDTSTNG 696
Db 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNYKVNYAVTKENTIINPSENGDTSTNG 720
Qy 697 IKKILIFSCKGYEIG 711
Db 721 IKKILIFSCKGYEIG 735

RESULT 12
US-09-848-909-11
; Sequence 11, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-11

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPELKOKGSNNRKKRSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPELKOKGSNNRKKRSTAGPTVPDRDN 180
Qy 181 DGIPLSLEVEGTVVVKNNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYDFEKVT 240
```

```
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTSPYDFEKT 240
Qy 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVAHSAFFDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYNNVLTPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYNNVLTPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 397 LNAQKQDASSTPTIMYNNQFLELEKTKQLRLDQVYGNIAATYNFENGRVVRVDTGSNWSEV 456
Db 421 LNAQKQDASSTPTIMYNNQFLELEKTKQLRLDQVYGNIAATYNFENGRVVRVDTGSNWSEV 480
Qy 457 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 516
Db 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 540
Qy 517 QYQKQDITFDFNFDOQTQNT:KNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 576
Db 541 QYQKQDITFDFNFDOQTQNT:KNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 600
Qy 577 NNIAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 637 DMLNISLRODQKTFIDFKKNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 696
Db 661 DMLNISLRODQKTFIDFKKNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Qy 697 IKKILIFSKKGVEIG 711
Db 721 IKKILIFSKKGVEIG 735

RESULT 13
US-09-848-909-12
; Sequence 12, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-12

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESQGLGYYFSDLNFPQAPMVVTSSTGDLSPSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLGYYFSDLNFPQAPMVVTSSTGDLSPSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDBYTATSDADNHVTWVDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKSDBYTATSDADNHVTWVDQEVINKASNSKIRLEKGLYQIKIY 120
```

```
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKRCSTASGPTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKRCSTASGPTVPDRN 180
Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTSPYDFEKT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTSPYDFEKT 240
Qy 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVAHSAFFDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYNNVLTPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYNNVLTPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 397 LNAQKQDASSTPTIMYNNQFLELEKTKQLRLDQVYGNIAATYNFENGRVVRVDTGSNWSEV 456
Db 421 LNAQKQDASSTPTIMYNNQFLELEKTKQLRLDQVYGNIAATYNFENGRVVRVDTGSNWSEV 480
Qy 457 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 516
Db 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 540
Qy 517 QYQKQDITFDFNFDOQTQNT:KNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 576
Db 541 QYQKQDITFDFNFDOQTQNT:KNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 600
Qy 577 NNIAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 637 DMLNISLRODQKTFIDFKKNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 696
Db 661 DMLNISLRODQKTFIDFKKNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Qy 697 IKKILIFSKKGVEIG 711
Db 721 IKKILIFSKKGVEIG 735

RESULT 14
US-09-848-909-13
; Sequence 13, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-13

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESQGLGYYFSDLNFPQAPMVVTSSTGDLSPSSELENIPSENQYF 60
```

```
Db 1 EVKQENRLNSESSESSQGLLYFSDLPQAPMVVTSSTGDLSPSSSELENIPESENQYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Qy 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPRDN 180
Qy 181 DGIPODSEVBEGYVDVQNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDSEKVT 240
Db 181 DGIPODSEVBEGYVDVQNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDSEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASPFIDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 397 LNAQKQDASPTITMNYNQFLEKTKQRLDQVYGNIAATYNFENGRVVRTGSKNWEV 456
Db 421 LNAQKQDASPTITMNYNQFLEKTKQRLDQVYGNIAATYNFENGRVVRTGSKNWEV 480
Qy 457 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGNEPNGL 516
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGNEPNGL 540
Qy 517 QYQKQDITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 576
Db 541 QYQKQDITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600
Qy 577 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 636
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 637 DMLNSSLRODQKTFIDFKYNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 696
Db 661 DMLNSSLRODQKTFIDFKYNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Qy 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735
```

RESULT 15

```
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-14
```

```
Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLLYFSDLPQAPMVVTSSTGDLSPSSSELENIPESENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYFSDLPQAPMVVTSSTGDLSPSSSELENIPESENQYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Qy 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPRDN 180
Qy 181 DGIPODSEVBEGYVDVQNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDSEKVT 240
Db 181 DGIPODSEVBEGYVDVQNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDSEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASPFIDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 397 LNAQKQDASPTITMNYNQFLEKTKQRLDQVYGNIAATYNFENGRVVRTGSKNWEV 456
Db 421 LNAQKQDASPTITMNYNQFLEKTKQRLDQVYGNIAATYNFENGRVVRTGSKNWEV 480
Qy 457 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGNEPNGL 516
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGNEPNGL 540
Qy 517 QYQKQDITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 576
Db 541 QYQKQDITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600
Qy 577 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 636
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 637 DMLNSSLRODQKTFIDFKYNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 696
Db 661 DMLNSSLRODQKTFIDFKYNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Qy 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735
```

Search completed: May 3, 2004, 20:02:16
Job time : 38.4429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 169.814 Seconds
(without alignments)
4086.665 Million cell updates/sec

Title: US-09-848-909A-16
Perfect score: 3648
Sequence: 1 EVKQENRLNSESQSL.....TSTNGIKLIPKKGVEIG 711

Scoring table: BLOSUM62
Gap0 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_Main:

```
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/2/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/2/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/2/paa/US107_COMB.pep.*
33: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	3648	100.0	711	1	PCT-US03-35733-16	Sequence 16, Appl
2	3648	100.0	711	23	US-09-848-909A-16	Sequence 16, Appl
3	3642	99.8	711	1	PCT-US03-35733-15	Sequence 15, Appl
4	3642	99.8	711	23	US-09-848-909A-15	Sequence 15, Appl
5	3635	99.6	711	1	PCT-US03-35733-14	Sequence 14, Appl
6	3635	99.6	711	23	US-09-848-909A-14	Sequence 14, Appl
7	3629	99.5	711	1	PCT-US03-35733-12	Sequence 12, Appl
8	3629	99.5	711	23	US-09-848-909A-12	Sequence 12, Appl
9	3626	99.4	735	1	PCT-US03-35733-13	Sequence 13, Appl
10	3626	99.4	735	23	US-09-848-909A-13	Sequence 13, Appl
11	3620	99.2	735	1	PCT-US03-35733-10	Sequence 10, Appl
12	3620	99.2	735	23	US-09-848-909A-10	Sequence 10, Appl
13	3614	99.1	735	1	PCT-US03-19786-4	Sequence 4, Appl
14	3614	99.1	735	23	US-09-848-909A-2	Sequence 2, Appl
15	3614	99.1	735	23	US-09-848-909A-2	Sequence 2, Appl
16	3613	99.0	735	1	PCT-US03-35733-8	Sequence 8, Appl
17	3613	99.0	735	1	PCT-US03-35733-9	Sequence 9, Appl
18	3613	99.0	735	23	US-09-848-909A-8	Sequence 8, Appl
19	3613	99.0	735	23	US-09-848-909A-8	Sequence 8, Appl
20	3609	98.9	735	1	PCT-US03-35733-7	Sequence 7, Appl
21	3609	98.9	735	1	PCT-US03-35733-23	Sequence 23, Appl
22	3609	98.9	735	23	US-09-848-909A-7	Sequence 7, Appl
23	3609	98.9	735	23	US-09-848-909A-23	Sequence 23, Appl
24	3608	98.9	735	1	PCT-US03-35733-4	Sequence 4, Appl
25	3608	98.9	735	1	PCT-US03-35733-6	Sequence 6, Appl
26	3608	98.9	735	1	PCT-US03-35733-18	Sequence 18, Appl
27	3608	98.9	735	23	US-09-848-909A-4	Sequence 4, Appl
28	3608	98.9	735	23	US-09-848-909A-6	Sequence 6, Appl
29	3608	98.9	735	23	US-09-848-909A-18	Sequence 18, Appl
30	3607	98.9	762	4	US-08-042-745-5	Sequence 5, Appl
31	3607	98.9	735	1	PCT-US03-35733-5	Sequence 5, Appl
32	3607	98.9	735	1	PCT-US03-35733-11	Sequence 11, Appl
33	3607	98.9	735	1	PCT-US03-35733-17	Sequence 17, Appl
34	3607	98.9	735	1	PCT-US03-35733-19	Sequence 19, Appl
35	3607	98.9	735	1	PCT-US03-35733-20	Sequence 20, Appl
36	3607	98.9	735	1	PCT-US03-35733-21	Sequence 21, Appl
37	3607	98.9	735	22	US-09-791-537-43735	Sequence 43735, A
38	3607	98.9	735	23	US-09-848-909A-5	Sequence 5, Appl
39	3607	98.9	735	23	US-09-848-909A-17	Sequence 17, Appl
40	3607	98.9	735	23	US-09-848-909A-19	Sequence 19, Appl
41	3607	98.9	735	23	US-09-848-909A-20	Sequence 20, Appl
42	3607	98.9	735	23	US-09-848-909A-21	Sequence 21, Appl
43	3607	98.9	735	23	US-09-848-909A-21	Sequence 21, Appl
44	3607	98.9	735	30	US-10-410-647-30	Sequence 30, Appl
45	3607	98.9	736	1	PCT-US01-14372A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
PCT-US03-35733-16
; Sequence 16, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-16

Query Match 100.0%; Score 3648; DB 1; Length 711;
Best Local Similarity 100.0%; Pred. No. 6.8e-287;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFAQPMVVTSSTTGDLSTPSSSELENIPISENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYGYFSDLNFAQPMVVTSSTTGDLSTPSSSELENIPISENOYF 60
QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVIVHVDMMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVIVHVDMMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 180
QY 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXWSTASDPYDFEKT 240
Db 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXWSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRVYNTGTAPIYVLPITSLV 360
Db 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRVYNTGTAPIYVLPITSLV 360
QY 361 LGKQNLATIKADENQSLQILAPNNYPSKNLAPIALNAQKADASTPTIMYNOFLELEK 420
Db 361 LGKQNLATIKADENQSLQILAPNNYPSKNLAPIALNAQKADASTPTIMYNOFLELEK 420
QY 421 TKQLRLDQVGNATYATNFENGRVVDVTSNWSSEVLPOIQETTARIIFNGKDLNLVERR 480
Db 421 TKQLRLDQVGNATYATNFENGRVVDVTSNWSSEVLPOIQETTARIIFNGKDLNLVERR 480
QY 481 IAAVNPSPLETTKPDMTLKEALKIAFGNPNENGLQYQKDIETFDNFQDQTSQNIKN 540
Db 481 IAAVNPSPLETTKPDMTLKEALKIAFGNPNENGLQYQKDIETFDNFQDQTSQNIKN 540
QY 541 QLAELNATNIYTVLDKIKLNAMNLIIRDKRHYDRNNIAGADESVVKEAREHVINST 600
Db 541 QLAELNATNIYTVLDKIKLNAMNLIIRDKRHYDRNNIAGADESVVKEAREHVINST 600
QY 601 EGGLELNDIKDILKILSGYIVEIEDTEGLKEVINDRYDMLNISLRLQDQKTFIDFKKYNDK 660
Db 601 EGGLELNDIKDILKILSGYIVEIEDTEGLKEVINDRYDMLNISLRLQDQKTFIDFKKYNDK 660
QY 661 LPYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711
Db 661 LPYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711
```

RESULT 2

```
US-09-848-909A-16
; Sequence 16, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-16
```

Query Match 100.0%; Score 3648; DB 23; Length 711;
Best Local Similarity 100.0%; Pred. No. 6.8e-287;

```
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFAQPMVVTSSTTGDLSTPSSSELENIPISENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYGYFSDLNFAQPMVVTSSTTGDLSTPSSSELENIPISENOYF 60
QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVIVHVDMMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVIVHVDMMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 180
QY 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXWSTASDPYDFEKT 240
Db 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXWSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRVYNTGTAPIYVLPITSLV 360
Db 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRVYNTGTAPIYVLPITSLV 360
QY 361 LGKQNLATIKADENQSLQILAPNNYPSKNLAPIALNAQKADASTPTIMYNOFLELEK 420
Db 361 LGKQNLATIKADENQSLQILAPNNYPSKNLAPIALNAQKADASTPTIMYNOFLELEK 420
QY 421 TKQLRLDQVGNATYATNFENGRVVDVTSNWSSEVLPOIQETTARIIFNGKDLNLVERR 480
Db 421 TKQLRLDQVGNATYATNFENGRVVDVTSNWSSEVLPOIQETTARIIFNGKDLNLVERR 480
QY 481 IAAVNPSPLETTKPDMTLKEALKIAFGNPNENGLQYQKDIETFDNFQDQTSQNIKN 540
Db 481 IAAVNPSPLETTKPDMTLKEALKIAFGNPNENGLQYQKDIETFDNFQDQTSQNIKN 540
QY 541 QLAELNATNIYTVLDKIKLNAMNLIIRDKRHYDRNNIAGADESVVKEAREHVINST 600
Db 541 QLAELNATNIYTVLDKIKLNAMNLIIRDKRHYDRNNIAGADESVVKEAREHVINST 600
QY 601 EGGLELNDIKDILKILSGYIVEIEDTEGLKEVINDRYDMLNISLRLQDQKTFIDFKKYNDK 660
Db 601 EGGLELNDIKDILKILSGYIVEIEDTEGLKEVINDRYDMLNISLRLQDQKTFIDFKKYNDK 660
QY 661 LPYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711
Db 661 LPYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711
```

RESULT 3

```
PCT-US03-35733-15
; Sequence 15, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-15
```

Query Match 99.8%; Score 3642; DB 1; Length 711;
Best Local Similarity 99.9%; Pred. No. 2.1e-286;

Matches 710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60

QY 61 QSAIWSGFTKVKKSDYFTFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFTKVKKSDYFTFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENTEKGLDFKLYWTSQNKKEVISSDNQLQPLKQKSSNSRKRSTSGAGTVPDRDN 180
DB 121 QRENTEKGLDFKLYWTSQNKKEVISSDNQLQPLKQKSSNSRKRSTSGAGTVPDRDN 180

QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASPYSDFEKVT 240
DB 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASPYSDFEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300

QY 301 SNSNSTVAIDHSLSLAGERTWAEITGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360
DB 301 SNSNSTVAIDHSLSLAGERTWAEITGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360

QY 361 LGKNOTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPIITMNYNQFLELEK 420
DB 361 LGKNOTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPIITMNYNQFLELEK 420

QY 421 TKQLRLDQVYGNATYATNFENGRVVDGTSNWSSEVLPOIQTETARIIFNGKDLNLVER 480
DB 421 TKQLRLDQVYGNATYATNFENGRVVDGTSNWSSEVLPOIQTETARIIFNGKDLNLVER 480

QY 481 IAAVNSPDPLETTKPDWTLKEALKIAGFNEPENGLOQYQKDIETEFDFNFQDQTSQNIKN 540
DB 481 IAAVNSPDPLETTKPDWTLKEALKIAGFNEPENGLOQYQKDIETEFDFNFQDQTSQNIKN 540

QY 541 QLAELNATNIYTVLDKIKLNAKXNLIIRDKRHYDRNNIAGADESVVKEAHREVINSST 600
DB 541 QLAELNATNIYTVLDKIKLNAKXNLIIRDKRHYDRNNIAGADESVVKEAHREVINSST 600

QY 601 EGLLLNIDKDIRKILSGYVEIETEGLEKEVINDRYDMLNSSLFQDQGTFTDFKKNYDK 660
DB 601 EGLLLNIDKDIRKILSGYVEIETEGLEKEVINDRYDMLNSSLFQDQGTFTDFKKNYDK 660

QY 661 LPLYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711
DB 661 LPLYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711

RESULT 4

US-09-848-909A-15

; Sequence 15, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Cellier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-15

Query Match

99.8%; Score 3642; DB 23; Length 711;

Best Local Similarity 99.9%; Pred. No. 2.1e-286;
Matches 710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60

QY 61 QSAIWSGFTKVKKSDYFTFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFTKVKKSDYFTFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENTEKGLDFKLYWTSQNKKEVISSDNQLQPLKQKSSNSRKRSTSGAGTVPDRDN 180
DB 121 QRENTEKGLDFKLYWTSQNKKEVISSDNQLQPLKQKSSNSRKRSTSGAGTVPDRDN 180

QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASPYSDFEKVT 240
DB 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASPYSDFEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300

QY 301 SNSNSTVAIDHSLSLAGERTWAEITGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360
DB 301 SNSNSTVAIDHSLSLAGERTWAEITGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360

QY 361 LGKNOTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPIITMNYNQFLELEK 420
DB 361 LGKNOTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPIITMNYNQFLELEK 420

QY 421 TKQLRLDQVYGNATYATNFENGRVVDGTSNWSSEVLPOIQTETARIIFNGKDLNLVER 480
DB 421 TKQLRLDQVYGNATYATNFENGRVVDGTSNWSSEVLPOIQTETARIIFNGKDLNLVER 480

QY 481 IAAVNSPDPLETTKPDWTLKEALKIAGFNEPENGLOQYQKDIETEFDFNFQDQTSQNIKN 540
DB 481 IAAVNSPDPLETTKPDWTLKEALKIAGFNEPENGLOQYQKDIETEFDFNFQDQTSQNIKN 540

QY 541 QLAELNATNIYTVLDKIKLNAKXNLIIRDKRHYDRNNIAGADESVVKEAHREVINSST 600
DB 541 QLAELNATNIYTVLDKIKLNAKXNLIIRDKRHYDRNNIAGADESVVKEAHREVINSST 600

QY 601 EGLLLNIDKDIRKILSGYVEIETEGLEKEVINDRYDMLNSSLFQDQGTFTDFKKNYDK 660
DB 601 EGLLLNIDKDIRKILSGYVEIETEGLEKEVINDRYDMLNSSLFQDQGTFTDFKKNYDK 660

QY 661 LPLYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711
DB 661 LPLYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711

RESULT 5

PCT-US03-35733-14

; Sequence 14, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-14

Query Match

99.6%; Score 3635; DB 1; Length 711;

Best Local Similarity 99.7%; Pred. No. 7.7e-286;
Matches 709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	EVKQENRLNESSSQGLLGYVFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Db	1	EVKQENRLNESSSQGLLGYVFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Qy	61	QSAIWSGFIKVKKSDBYTATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY	120
Db	61	QSAIWSGFIKVKKSDBYTATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY	120
Qy	121	QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKKRSTASAGTVPDRDN	180
Db	121	QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKKRSTASAGTVPDRDN	180
Qy	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXSTASDPYDFEKT	240
Db	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXSTASDPYDFEKT	240
Qy	241	GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTRTHT	300
Db	241	GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTRTHT	300
Qy	301	SNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTSLV	360
Db	301	SNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTSLV	360
Qy	361	LGKQTLATIKADENQSLQILAPNNYPSKNLAPIALNAQKQDASSTPTIMYNOFLELEK	420
Db	361	LGKQTLATIKADENQSLQILAPNNYPSKNLAPIALNAQDASSTPTIMYNOFLELEK	420
Qy	421	TKQLRLDTQVYGNIAATYFNGRVRVDTGSKNWEVLPOIQTETARIIFNGKDLNLEVER	480
Db	421	TKQLRLDTQVYGNIAATYFNGRVRVDTGSKNWEVLPOIQTETARIIFNGKDLNLEVER	480
Qy	481	IAAVNSDPLETTKPDMTLKEALKIAGFNEPNNGNLOYGKOITEFDNFDOQTSONIKN	540
Db	481	IAAVNSDPLETTKPDMTLKEALKIAGFNEPNNGNLOYGKOITEFDNFDOQTSONIKN	540
Qy	541	QLAELNATNIYTVLDKIKLNAKXNIIIRDKRPHYDRNNIAGADESVVKEAREVINSST	600
Db	541	QLAELNATNIYTVLDKIKLNAKXNIIIRDKRPHYDRNNIAGADESVVKEAREVINSST	600
Qy	601	EGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRYDMLNIISSLRQDGKTFIDFKKYNDK	660
Db	601	EGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRYDMLNIISSLRQDGKTFIDFKKYNDK	660
Qy	661	LPLYISPNYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG	711
Db	661	LPLYISPNYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG	711

RESULT 6

US-09-848-909A-14
; Sequence 14, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-14

RESULT 7

PCT-US03-35733-12
; Sequence 12, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-12

Query Match 99.6%; Score 3635; DB 23; Length 711;
Best Local Similarity 99.7%; Pred. No. 7.7e-286;
Matches 709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	EVKQENRLNESSSQGLLGYVFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Db	1	EVKQENRLNESSSQGLLGYVFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Qy	61	QSAIWSGFIKVKKSDBYTATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY	120
Db	61	QSAIWSGFIKVKKSDBYTATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY	120
Qy	121	QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKKRSTASAGTVPDRDN	180
Db	121	QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKKRSTASAGTVPDRDN	180
Qy	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXSTASDPYDFEKT	240
Db	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXSTASDPYDFEKT	240
Qy	241	GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTRTHT	300
Db	241	GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTRTHT	300
Qy	301	SNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTSLV	360
Db	301	SNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTSLV	360
Qy	361	LGKQTLATIKADENQSLQILAPNNYPSKNLAPIALNAQKQDASSTPTIMYNOFLELEK	420
Db	361	LGKQTLATIKADENQSLQILAPNNYPSKNLAPIALNAQDASSTPTIMYNOFLELEK	420
Qy	421	TKQLRLDTQVYGNIAATYFNGRVRVDTGSKNWEVLPOIQTETARIIFNGKDLNLEVER	480
Db	421	TKQLRLDTQVYGNIAATYFNGRVRVDTGSKNWEVLPOIQTETARIIFNGKDLNLEVER	480
Qy	481	IAAVNSDPLETTKPDMTLKEALKIAGFNEPNNGNLOYGKOITEFDNFDOQTSONIKN	540
Db	481	IAAVNSDPLETTKPDMTLKEALKIAGFNEPNNGNLOYGKOITEFDNFDOQTSONIKN	540
Qy	541	QLAELNATNIYTVLDKIKLNAKXNIIIRDKRPHYDRNNIAGADESVVKEAREVINSST	600
Db	541	QLAELNATNIYTVLDKIKLNAKXNIIIRDKRPHYDRNNIAGADESVVKEAREVINSST	600
Qy	601	EGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRYDMLNIISSLRQDGKTFIDFKKYNDK	660
Db	601	EGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRYDMLNIISSLRQDGKTFIDFKKYNDK	660
Qy	661	LPLYISPNYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG	711
Db	661	LPLYISPNYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG	711

Query Match 99.5%; Score 3629; DB 1; Length 711;
 Best Local Similarity 99.6%; Pred. No. 2.4e-285;
 Matches 708; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQYF 60
 DB 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKRLYQIKIY 120
 DB 61 QSAIWSGFIKVKKSDDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKRKSTASGTPVDRDN 180
 DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKRKSTASGTPVDRDN 180

QY 181 DGIPDSLEVEGYTVDYKNGRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDPYDFEKT 240
 DB 181 DGIPDSLEVEGYTVDYKNGRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
 DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360
 DB 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360

QY 361 LGKNOTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPTIMYNOFLELEK 420
 DB 361 LGKNOTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPTIMYNOFLELEK 420

QY 421 TKQLRLDTPQVYGNATYNFENGRVVDVTGSNWSVLPOIQTETARIIFNGKDLNVER 480
 DB 421 TKQLRLDTPQVYGNATYNFENGRVVDVTGSNWSVLPOIQTETARIIFNGKDLNVER 480

QY 481 IAAVNSPDLPTTKPDMTLKEALKIAGFNEPENGNYQOGKDIETFDNFDOOTSQNIKN 540
 DB 481 IAAVNSPDLPTTKPDMTLKEALKIAGFNEPENGNYQOGKDIETFDNFDOOTSQNIKN 540

QY 541 QLAELNATNIYTVLDKIKLNKXNILLIRDKRPHYDRNNIAVGADESUVKEAHEVINSST 600
 DB 541 QLAELNATNIYTVLDKIKLNKXNILLIRDKRPHYDRNNIAVGADESUVKEAHEVINSST 600

QY 601 EGGLELNDKIDKILSGYIIEIETDGLKEVINDRYDMLNSSLRODGKTFIDFKKYNK 660
 DB 601 EGGLELNDKIDKILSGYIIEIETDGLKEVINDRYDMLNSSLRODGKTFIDFKKYNK 660

QY 661 LPLVSNPNYKXNVYAVTKENTIINPSENGDTSTNGIKKILIFSCKGYEIG 711
 DB 661 LPLVSNPNYKXNVYAVTKENTIINPSENGDTSTNGIKKILIFSCKGYEIG 711

RESULT 8
 US-09-848-909A-12
 ; Sequence 12, Application US/09848909A
 ; GENERAL INFORMATION:
 ; APPLICANT: Collier, R. John
 ; TITLE OF INVENTION: Compounds and Methods for the Treatment
 ; TITLE OF INVENTION: and Prevention of Bacterial Infection
 ; FILE REFERENCE: 00742/060002
 ; CURRENT APPLICATION NUMBER: US/09/848,909A
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: US 60/201,800
 ; PRIOR FILING DATE: 2000-05-04
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 711
 ; TYPE: PRT
 ; ORGANISM: Bacillus anthracis
 US-09-848-909A-12

RESULT 9
 PCT-US03-35733-13
 ; Sequence 13, Application PC/TUS0335733
 ; GENERAL INFORMATION:
 ; APPLICANT: President and Fellows of Harvard College et al.
 ; TITLE OF INVENTION: Compounds and Methods for the Treatment
 ; TITLE OF INVENTION: and Prevention of Bacterial Infection
 ; FILE REFERENCE: 00742/072003
 ; CURRENT APPLICATION NUMBER: PCT/US03/35733
 ; CURRENT FILING DATE: 2003-10-10
 ; PRIOR APPLICATION NUMBER: US 60/424,987
 ; PRIOR FILING DATE: 2002-11-08
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 735
 ; TYPE: PRT
 ; ORGANISM: Bacillus anthracis
 PCT-US03-35733-13

Query Match 99.4%; Score 3626; DB 1; Length 735;
Best Local Similarity 96.7%; Pred. No. 4.4e-285;
Matches 711; Conservative 0; Mismatches 0; Indels 24; Gaps 1;
US-09-848-909A-13

QY 1 EVKQNRLLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
DB 1 EVKQNRLLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLQPLKQKSSNSRKRSTSGPTVPDRN 180
DB 121 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLQPLKQKSSNSRKRSTSGPTVPDRN 180
QY 181 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLQPLKQKSSNSRKRSTSGPTVPDRN 240
DB 181 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLQPLKQKSSNSRKRSTSGPTVPDRN 240
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 397 LNAQKASSTPTIMYVNFQLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 456
DB 421 LNAQKASSTPTIMYVNFQLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480
QY 457 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAGFNEPENG 516
DB 481 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAGFNEPENG 540
QY 517 QYQKIDITEFDNFDDQSTQNKQLAELNATNIYVLDKIKLNKAKNILLIRKRFHYDR 576
DB 541 QYQKIDITEFDNFDDQSTQNKQLAELNATNIYVLDKIKLNKAKNILLIRKRFHYDR 600
QY 577 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 636
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 637 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 696
DB 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
DB 721 IKKILIFSKKGYEIG 735

RESULT 10

US-09-848-909A-13
Sequence 13, Application US/09848909A
GENERAL INFORMATION:
APPLICANT: Sellman, R. John
TITLE OF INVENTION: Compounds and Methods for the Treatment
of Bacterial Infection
FILE REFERENCE: 00742/060002
CURRENT APPLICATION NUMBER: US/09/848,909A
PRIORITY FILING DATE: 2001-05-04
PRIORITY FILING DATE: 2000-05-04
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 4.0

RESULT 11

PCT-US03-35733-10
Sequence 10, Application PC/TUS0335733
GENERAL INFORMATION:
APPLICANT: President and Fellows of Harvard College et al.
TITLE OF INVENTION: Compounds and Methods for the Treatment
of Bacterial Infection
FILE REFERENCE: 00742/072003
CURRENT APPLICATION NUMBER: PCT/US03/35733
CURRENT FILING DATE: 2003-10-10

SEQ ID NO 13
LENGTH: 735
TYPE: PRT
ORGANISM: Bacillus anthracis
US-09-848-909A-13

Query Match 99.4%; Score 3626; DB 23; Length 735;
Best Local Similarity 96.7%; Pred. No. 4.4e-285;
Matches 711; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 EVKQNRLLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
DB 1 EVKQNRLLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLQPLKQKSSNSRKRSTSGPTVPDRN 180
DB 121 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLQPLKQKSSNSRKRSTSGPTVPDRN 180
QY 181 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLQPLKQKSSNSRKRSTSGPTVPDRN 240
DB 181 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLQPLKQKSSNSRKRSTSGPTVPDRN 240
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 397 LNAQKASSTPTIMYVNFQLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 456
DB 421 LNAQKASSTPTIMYVNFQLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480
QY 457 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAGFNEPENG 516
DB 481 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAGFNEPENG 540
QY 517 QYQKIDITEFDNFDDQSTQNKQLAELNATNIYVLDKIKLNKAKNILLIRKRFHYDR 576
DB 541 QYQKIDITEFDNFDDQSTQNKQLAELNATNIYVLDKIKLNKAKNILLIRKRFHYDR 600
QY 577 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 636
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 637 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 696
DB 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
QY 697 IKKILIFSKKGYEIG 711
DB 721 IKKILIFSKKGYEIG 735

; PRIOR APPLICATION NUMBER: US 60/424,987
 ; PRIOR FILING DATE: 2002-11-08
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 735
 ; TYPE: PRT
 ; ORGANISM: Bacillus anthracis
 PCT-US03-35733-10

Query Match 99.2%; Score 3620; DB 1; Length 735;
 Best Local Similarity 96.6%; Pred. No. 1.4e-284;
 Matches 710; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY	1	EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF	60
Db	1	EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF	60
QY	61	QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIY	120
Db	61	QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIY	120
QY	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPTVPDRDN	180
Db	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPTVPDRDN	180
QY	181	DGIPDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYDFEKT	240
Db	181	DGIPDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYDFEKT	240
QY	241	GRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSTHT	300
Db	241	GRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSTHT	300
QY	301	-----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	336
Db	301	SEVHGNAEVHASFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	337	NANIRYVNTGTAPIYINVLPFTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA	396
Db	361	NANIRYVNTGTAPIYINVLPFTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA	420
QY	397	LNAQKXASSTPTIMNYNQFLEKTKQLRLDQVYGNIAATYNFENGVRVDTGNSWSEV	456
Db	421	LNAQKXASSTPTIMNYNQFLEKTKQLRLDQVYGNIAATYNFENGVRVDTGNSWSEV	480
QY	457	LPOIQTETARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIAGFNEPNGNL	516
Db	481	LPOIQTETARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIAGFNEPNGNL	540
QY	517	QYQKDIITEFDNFDDQSTSONIKNQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR	576
Db	541	QYQKDIITEFDNFDDQSTSONIKNQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR	600
QY	577	NNIAYGADSVVKEAHRVINSSTEGLLNIDKDKIRKILSGYIIEIDTEGLKEVINDRY	636
Db	601	NNIAYGADSVVKEAHRVINSSTEGLLNIDKDKIRKILSGYIIEIDTEGLKEVINDRY	660
QY	637	DMNLNSSLRQDGKTFIDFKKYNKPLVIYISNPKNYKVVYAVTKENTIINPSENGDTSTNG	696
Db	661	DMNLNSSLRQDGKTFIDFKKYNKPLVIYISNPKNYKVVYAVTKENTIINPSENGDTSTNG	720
QY	697	IKKILIFSKKGYEIG 711	
Db	721	IKKILIFSKKGYEIG 735	

RESULT 12

US-09-848-909a-10

; Sequence 10, Application US/09848909a
 ; GENERAL INFORMATION:
 ; APPLICANT: Collier, R. John
 ; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment
 ; TITLE OF INVENTION: and Prevention of Bacterial Infection
 ; FILE REFERENCE: 00742/060002
 ; CURRENT APPLICATION NUMBER: US/09/848,909A
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: US 60/201,800
 ; PRIOR FILING DATE: 2000-05-04
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 735
 ; TYPE: PRT
 ; ORGANISM: Bacillus anthracis
 US-09-848-909a-10

Query Match 99.2%; Score 3620; DB 23; Length 735;
 Best Local Similarity 96.6%; Pred. No. 1.4e-284;
 Matches 710; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY	1	EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF	60
Db	1	EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF	60
QY	61	QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIY	120
Db	61	QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIY	120
QY	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPTVPDRDN	180
Db	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPTVPDRDN	180
QY	181	DGIPDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYDFEKT	240
Db	181	DGIPDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYDFEKT	240
QY	241	GRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSTHT	300
Db	241	GRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSTHT	300
QY	301	-----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	336
Db	301	SEVHGNAEVHASFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	337	NANIRYVNTGTAPIYINVLPFTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA	396
Db	361	NANIRYVNTGTAPIYINVLPFTSLVLGKNQTLATIKADENQLSQILAPNNYYPKSLAPIA	420
QY	397	LNAQKXASSTPTIMNYNQFLEKTKQLRLDQVYGNIAATYNFENGVRVDTGNSWSEV	456
Db	421	LNAQKXASSTPTIMNYNQFLEKTKQLRLDQVYGNIAATYNFENGVRVDTGNSWSEV	480
QY	457	LPOIQTETARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIAGFNEPNGNL	516
Db	481	LPOIQTETARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIAGFNEPNGNL	540
QY	517	QYQKDIITEFDNFDDQSTSONIKNQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR	576
Db	541	QYQKDIITEFDNFDDQSTSONIKNQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR	600
QY	577	NNIAYGADSVVKEAHRVINSSTEGLLNIDKDKIRKILSGYIIEIDTEGLKEVINDRY	636
Db	601	NNIAYGADSVVKEAHRVINSSTEGLLNIDKDKIRKILSGYIIEIDTEGLKEVINDRY	660
QY	637	DMNLNSSLRQDGKTFIDFKKYNKPLVIYISNPKNYKVVYAVTKENTIINPSENGDTSTNG	696
Db	661	DMNLNSSLRQDGKTFIDFKKYNKPLVIYISNPKNYKVVYAVTKENTIINPSENGDTSTNG	720
QY	697	IKKILIFSKKGYEIG 711	
Db	721	IKKILIFSKKGYEIG 735	

RESULT 13

PCT-US03-19786-4
; Sequence 4, Application PC/TUS0319786
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Modified Shine Dalgarno Sequences and Methods of Use Thereof
; FILE REFERENCE: PV595PCT
; CURRENT APPLICATION NUMBER: PCT/US03/19786
; CURRENT FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Mature PA sequence including an ETB signal sequence
PCT-US03-19786-4

Query Match 99.2%; Score 3617; DB 1; Length 735;
Best Local Similarity 96.5%; Pred. No. 2.4e-284;
Matches 709; Conservative 1; Mismatches 1; Indels 24; Gaps 1;

QY 1 EVKQENRLNESSSOGLLYGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSOGLLYGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTMVDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTMVDQEVINKASNSNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKRSTASGTPVDPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKRSTASGTPVDPDRN 180

QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240
Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETTSKNTSTSRTH 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETTSKNTSTSRTH 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVAHFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 337 NANIRYNTGTAPIYVLPPTSLVLGKNOTLATIKADENQLSOLAPNNYPSKNLAPIA 396
Db 361 NANIRYNTGTAPIYVLPPTSLVLGKNOTLATIKADENQLSOLAPNNYPSKNLAPIA 420

QY 397 LNAQKQASSTPTIMYNOQFLEKTKQLRLDQVYGNIAIYFNENGRVVDVTSNWSV 456
Db 421 LNAQKQASSTPTIMYNOQFLEKTKQLRLDQVYGNIAIYFNENGRVVDVTSNWSV 480

QY 457 LPQIQTETARIIFNGKDLNVERIAAANPSDPLETTKPDMLKEALKIAGFNEPENG 516
Db 481 LPQIQTETARIIFNGKDLNVERIAAANPSDPLETTKPDMLKEALKIAGFNEPENG 540

QY 517 QYQKDIETFDNFDOQTSONIKQKLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 576
Db 541 QYQKDIETFDNFDOQTSONIKQKLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600

QY 577 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIYVEIEDTEGLKEVINDRY 636
Db 601 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIYVEIEDTEGLKEVINDRY 660

QY 637 DMLNLSLRQDKGTFFDKKYNKDLPLIYISNPNKVNYYAVTKENTIIINSENGDTSTNG 696
Db 661 DMLNLSLRQDKGTFFDKKYNKDLPLIYISNPNKVNYYAVTKENTIIINSENGDTSTNG 720

QY 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735

RESULT 14
PCT-US03-35733-2
; Sequence 2, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-2

Query Match 99.1%; Score 3614; DB 1; Length 735;
Best Local Similarity 96.5%; Pred. No. 4.2e-284;
Matches 709; Conservative 0; Mismatches 2; Indels 24; Gaps 1;

QY 1 EVKQENRLNESSSOGLLYGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSOGLLYGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTMVDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTMVDQEVINKASNSNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKRSTASGTPVDPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKRSTASGTPVDPDRN 180

QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240
Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETTSKNTSTSRTH 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETTSKNTSTSRTH 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVAHFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 337 NANIRYNTGTAPIYVLPPTSLVLGKNOTLATIKADENQLSOLAPNNYPSKNLAPIA 396
Db 361 NANIRYNTGTAPIYVLPPTSLVLGKNOTLATIKADENQLSOLAPNNYPSKNLAPIA 420

QY 397 LNAQKQASSTPTIMYNOQFLEKTKQLRLDQVYGNIAIYFNENGRVVDVTSNWSV 456
Db 421 LNAQKQASSTPTIMYNOQFLEKTKQLRLDQVYGNIAIYFNENGRVVDVTSNWSV 480

QY 457 LPQIQTETARIIFNGKDLNVERIAAANPSDPLETTKPDMLKEALKIAGFNEPENG 516
Db 481 LPQIQTETARIIFNGKDLNVERIAAANPSDPLETTKPDMLKEALKIAGFNEPENG 540

QY 517 QYQKDIETFDNFDOQTSONIKQKLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 576
Db 541 QYQKDIETFDNFDOQTSONIKQKLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600

QY 577 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIYVEIEDTEGLKEVINDRY 636
Db 601 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIYVEIEDTEGLKEVINDRY 660

QY 637 DMLNLSLRQDKGTFFDKKYNKDLPLIYISNPNKVNYYAVTKENTIIINSENGDTSTNG 696
Db 661 DMLNLSLRQDKGTFFDKKYNKDLPLIYISNPNKVNYYAVTKENTIIINSENGDTSTNG 720

Db 601 NNTAVGADESVMKEAHREVINSSTEGLLNIDKIDIRKILSGYIVEIETDTEGLKEVINDRY 660
Qy 637 DMLNSSLRODQGTFFIDFKKYNDKLPYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 696
Db 661 DMLNSSLRODQGTFFIDFKKYNDKLPYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Qy 697 IKKILIFSCKGYEIG 711
Db 721 IKKILIFSCKGYEIG 735

Search completed: May 3, 2004, 19:57:47
Job time : 172.814 secs

RESULT 15

US-09-848-909A-2
; Sequence 2, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-2

Query Match 99.1%; Score 3614; DB 23; Length 735;
Best Local Similarity 96.5%; Pred No. 4.2e-284;
Matches 709; Conservative 0; Mismatches 2; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFFDLNTPQAPMVVTSSTTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFFDLNTPQAPMVVTSSTTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWGGFTKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKQY 120
Db 61 QSAIWGGFTKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKQY 120
Qy 121 QRENPTKGLDFKLYWTDSCNKEVISSDNQLQPELKQSSNSRKRKSTAGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSCNKEVISSDNQLQPELKQSSNSRKRKSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDPEKVT 240
Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDPEKVT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVEHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPYVNLPTTSLVLGKNQOTLATIKADENOLSOILAPNNYPSKNLAPTA 396
Db 361 NANIRYVNTGTAPYVNLPTTSLVLGKNQOTLATIKADENOLSOILAPNNYPSKNLAPTA 420
Qy 397 LNAQKASSTPIITNNYNQFLEKTKQLRLDTPQVYGNLATYNFENGVRVDTGNSWSEV 456
Db 421 LNAQDDFSPITNNYNQFLEKTKQLRLDTPQVYGNLATYNFENGVRVDTGNSWSEV 480
Qy 457 LPQIQETARIIFNGKDLNVERRIIAVNPSPDLETTPKDMTLKEALAFGENEENGNL 516
Db 481 LPQIQETARIIFNGKDLNVERRIIAVNPSPDLETTPKDMTLKEALAFGENEENGNL 540
Qy 517 QYQGKDITEFDNFDQTSQNIKNQLAELNATNIYVLDXIKLNAKNILIRKRFHYDR 576
Db 541 QYQGKDITEFDNFDQTSQNIKNQLAELNATNIYVLDXIKLNAKNILIRKRFHYDR 600
Qy 577 NNTAVGADESVMKEAHREVINSSTEGLLNIDKIDIRKILSGYIVEIETDTEGLKEVINDRY 636

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 12.6582 Seconds

(without alignments)

5403.004 Million cell updates/sec

Title: US-09-848-909A-16

Perfect score: 3648

Sequence: 1 EVQENRLNESSSQGLL.....TSTNGIKKILFISKKGYEIG 711

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3607	98.9	764	2 I39334	protective antigen
2	849	23.3	875	2 I40862	iota toxin compone
3	235.5	6.5	192	2 I39933	cryptic protein -
4	235.5	6.5	204	2 G5104	hypothetical prote
5	200	5.5	4888	2 F82885	hypothetical prote
6	197	5.4	1302	1 JC6009	surface-located me
7	192.5	5.3	2529	2 B64635	toxin-like outer m
8	190.5	5.2	2401	2 T28676	rhopty protein -
9	190	5.2	752	2 G90599	hypothetical prote
10	188	5.2	1658	2 S5101	hypothetical prote
11	187	5.1	2269	2 T28677	rhopty protein -
12	187	5.1	6713	2 B89921	hypothetical prote
13	181	5.0	4919	2 T31105	hypothetical prote
14	180	4.9	1227	2 C97033	uncharacterized pr
15	179.5	4.9	624	2 PC6003	surface membrane p
16	179.5	4.9	1365	2 C90822	lmp1 protein - Myc
17	178	4.9	1272	2 C90593	hypothetical prote
18	177.5	4.9	1072	2 A88827	hypothetical prote
19	177	4.9	1839	2 T18372	repeat organellar
20	175	4.8	3216	2 C90538	hypothetical prote
21	175	4.8	3724	2 T18427	hypothetical prote
22	174.5	4.8	821	2 S67087	hypothetical prote
23	174.5	4.8	1115	2 T41342	probable coiled-co
24	173.5	4.8	1125	2 E90598	membrane nuclease,
25	173.5	4.8	1639	2 S03603	major merozoite su
26	173.5	4.8	2399	2 H78879	toxin-like outer m
27	173	4.7	769	2 F89870	serine proteinase
28	171.5	4.7	1208	2 T30668	coiled coil protei
29	171	4.7	1315	2 T28679	fibrinogen-binding

30	171	4.7	1553	2 T18502	hypothetical prote
31	168.5	4.6	786	2 T18469	hypothetical prote
32	168	4.6	1612	2 AB1347	probable peptidogl
33	166.5	4.6	2526	2 T20531	hypothetical prote
34	166	4.6	1837	2 T41023	probable nuclear p
35	166	4.6	5005	2 F82884	hypothetical prote
36	165.5	4.5	1802	2 S52611	TyB protein - yeas
37	165	4.5	1308	2 E71622	probable membrane
38	165	4.5	2013	2 AD1129	probable peptidogl
39	164	4.5	2910	2 T28156	DNA-directed RNA p
40	163	4.5	1127	2 T28317	ORF MSV156 hypothe
41	163	4.5	1465	2 S31262	TyB protein - yeas
42	163	4.5	1631	1 SAZQK1	major merozoite su
43	163	4.5	1635	2 AI0452	hemolysin [importe
44	163	4.5	1803	2 S56894	TyB protein - yeas
45	162.5	4.5	1024	2 T30868	RhoA-binding prote

ALIGNMENTS

RESULT 1

I39934

protective antigen precursor - Bacillus anthracis plasmid

C:Species: Bacillus anthracis

C>Date: 19-Jul-1996 #sequence revision 19-Jul-1996 #text_change 01-Dec-2000

C:Accession: I39934; S69160; F59104

R:Melkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.

Gene 69, 287-300, 1998

A>Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr.

A:Reference number: I39933; MUID:89172073; PMID:3148491

A:Accession: I39934

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-764 <RES>

A:Cross-references: GB:M2589; NID:9143280; PIDN:AAA22637.1; PID:9143282

J:Friedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klumper, K.R.; Birch, N.P.; Loh, Y.P.

Arch. Biochem. Biophys. 316, 5-13, 1995

A>Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SP

A:Reference number: S69160; MUID:95142670; PMID:7840657

A:Accession: S69160

A:Molecule type: protein

A:Residues: 197-202 <FRI>

R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle

J. Bacteriol. 181, 6509-6515, 1999

A>Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbor

A:Reference number: A59091; MUID:99445483; PMID:10515943

A:Accession: F59104

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-313, 'Q', 315-764 <OKI>

A:Cross-references: GB:AF065404; NID:G4894216; PIDN:AAD32414.1; PID:G4894326

A:Experimental source: strain Sterne

A>Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid

C:Genetics:

A:Gene: pXO1-110

A:Genome: plasmid

C:Function:

A:Description: three component exotoxin; protective antigen binds to receptors on the s

Y active components edema factor or lethal factor; the complex is internalized by recep

C:Keywords: exotoxin

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-156/Domain: propeptide #status predicted <PRO>

F:197-202/Product: protective antigen #status experimental <MAT>

Query Match 98.9%; Score 3607; DB 2; Length 764;

Best Local Similarity 96.3%; Pred. No. 5.2e-173;

Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVQENRLNESSSQGLLGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60

Db 30 EVQENRLNESSSQGLLGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 89

QY 61 QSAIWSGFIKVKKSDXYTATFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 90 QSAIWSGFIKVKKSDXYTATFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 149
QY 121 QRENTKGLDFKLYWTDSONKKEVYSSDNQLQLPELKOKSNRKRKSTAGPTVPDRDN 180
DB 150 QRENTKGLDFKLYWTDSONKKEVYSSDNQLQLPELKOKSNRKRKSTAGPTVPDRDN 209
QY 181 DGIPOSLEVEGYTVVQKRTFLSPWISNIEHKGKLTYSKSPKSWSTASDPYDFEKT 240
DB 210 DGIPOSLEVEGYTVVQKRTFLSPWISNIEHKGKLTYSKSPKSWSTASDPYDFEKT 269
QY 241 GRIDKXVSPKSHPIVAAYPIVHVDMENIILSKNEDQSTQNTDSTRTISKNTSTSRHT 300
DB 270 GRIDKXVSPKSHPIVAAYPIVHVDMENIILSKNEDQSTQNTDSTRTISKNTSTSRHT 329
QY 301 -----SNSNSTVAIDHSLSLAGERTWAGTMTGLTADTARL 336
DB 330 SEVHGAEVHAFDIDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTMTGLTADTARL 389
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYYPKSNLAPIA 396
DB 390 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYYPKSNLAPIA 449
QY 397 LNAQKDASTPTIMYQNFLEKTKOLELTDQVYGNATYVNGRVRVDTGSKNSEV 456
DB 450 LNAQKDASTPTIMYQNFLEKTKOLELTDQVYGNATYVNGRVRVDTGSKNSEV 509
QY 457 LPOIQTETARIIFNGKDLNVERIAAANVPSDPLETTKPDMLKEALKIARFGNEPKNL 516
DB 510 LPOIQTETARIIFNGKDLNVERIAAANVPSDPLETTKPDMLKEALKIARFGNEPKNL 569
QY 517 QYQKDIITFDNFDOOTQNTKNQLAELNATYVTLVDKIKNAKMLIARDKRFHYDR 576
DB 570 QYQKDIITFDNFDOOTQNTKNQLAELNATYVTLVDKIKNAKMLIARDKRFHYDR 629
QY 577 NNIAVGADESVMKEAAREVINSSTEGLLNIDKIRKILSGYVIEIETEGLEKEVINDRY 636
DB 630 NNIAVGADESVMKEAAREVINSSTEGLLNIDKIRKILSGYVIEIETEGLEKEVINDRY 689
QY 637 DMLNISLQDQKTFIDFKYNDKPLIYSNPNYKVVAVTKENTIINPSENGDTSTNG 696
DB 690 DMLNISLQDQKTFIDFKYNDKPLIYSNPNYKVVAVTKENTIINPSENGDTSTNG 749
QY 697 IKKILFSSKGYEIG 711
DB 750 IKKILFSSKGYEIG 764

RESULT 2
I40862
iota toxin component Ib - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Oct-1999
C:Accession: I40862; S42774
R:Perelle, S.; Gibert, M.; Boquet, P.; Popoff, M.R.
Infect. Immun. 61, 5147-5156, 1993
A:Title: Characterization of Clostridium perfringens iota-toxin genes and expression in
A:Reference number: I40861; MUID:94041637; PMID:8225592
A:Accession: I40862
A:Status: preliminary; translated from GB/EMBL/DDSI
A:Molecule type: DNA
A:Residues: 1-875 <RES>
A:Cross-references: EMBL:X73562; NID:g929031; PIDN:CAA51960.1; PID:g414655

Query Match 23.3%; Score 849; DB 2; Length 875;
Best Local Similarity 31.3%; Pred. No. 5, 8e-35;
Matches 251; Conservative 128; Mismatches 268; Indels 154; Gaps 33;

QY 1 EVQENRLLNESSSQGLLGYYFSDLNFOAPMVVTSSTGDLISIPSSLENIPE-NQY 59
DB 35 DTNQEKEITTENTILSSNGLMGYYFADEHFKDLBMAPIKNGDLKFEKKVDTLTDNSS 94

QY 60 FOSAIWSGFIKVKKSDXYTATFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 119
DB 95 IKSIRWTGRIIPSEDEGEYILSTR-NDVLQINAKGDIKAK---TLKVMKKQAYNIRIE 150
QY 120 YQREN-----PTEKGLDFKLYWTDSONKKEVYSSDNQLQLPELKOKSNRKRKST 170
DB 151 IQDKNIGSIDNLSVP-----KLYW-ELUNGKRTWIPENLFRDYISKIDEND----- 195
QY 171 AGPTVP-----DRNDGIGIPDSLEVEGYTVVQKRTFLSPWISNHT 211
DB 196 --PFIENNFFDVRFPFAAWEDEDLTDNDNIPDAVEKNGYTI---KDSIAVKNDSPA 249
QY 212 EKUGLTKYKSSPKKSTASDPYDFEKTGRIDKXVSPKSHPIVAAYPIVHVDMENIIL 271
DB 250 E-QYKYYVSSVLESNTAGDPYDYQKASGSDKAIKLEARDPLVAAYPVGVGMENTLI 308
QY 272 SKNE-----DOSTQNTDSERTI-----SKNTSTSRHTSNSNSTVA 309
DB 309 STNEHASSDQKTVSRATNTSKTDANTVGVSIAGYQNGFTGNTIITSYHT--TDNSTAV 366
QY 310 IDSLSLAGERTWAGTMTGLTADTARLNAIRYVNTGTAPIYVNLPTTSLVLGKQTLAT 369
DB 367 QDSN---GB-SWNTGLSINKGESAYINANVRVYNTGTAPMYKVTPTTNLVL-DGETLAT 420
QY 370 IKADENQLSQILAPNNYYPKSNLAPIALNAQKDASTPTIMYQNFLEKTKOLELTD 429
DB 421 IKAQDNQIGNLSPNETYFKKGLSPLALNTMDFNARLIPINYDQLKLDGSKQIKLETT 480
QY 430 QYVGNATYVNFENGVRVDTGSKNSEVLPQIQTETARIIFNGKDLNVERIAAANVPSD 489
DB 481 QVSGNYGTKN-SQOQI-ITEGNSWNYISQIDSVSASIIID-TGSQTFERRVAAKEQGNP 537
QY 490 LETTKPDMLKEALKIARFGNEPKNLQY-QGKDIET--PDFNFDQOTQNTKNQLAELN 546
DB 538 EDKT-PEITIGEAIKKAFSAATK-NGELLYFNGIPIDESCVELIPDNTSEIKEQLKYLD 595
QY 547 ATNIYVTLVDKIKNAKMLIARDKRF--HYDR-NNIAVGADESVMKEAAREVINSSTEG 603
DB 596 DKKIYV---KLERGMNILLKVPSTFTNPFDEYNFP--ASWSNIDTKNQDGLQSVANKL 649
QY 604 -----LLNDKDIKILSGY-----IVEIEDTEGLEKEVINDRYDMLNIS-- 642
DB 650 SGETKIIIPMSKULPKRYVYFSGYKDPSTNSITVNIKSKEQKTYLVPKDYTFESYE 709
QY 643 -----SLRQDQKTFIDFKYND--KLPLYISNPNYKVV-----NVVA 676
DB 710 FETTKDSSDIETLTSSGVIPLDNLISITELNTPBILKEPEIKVPDQIBLDHAKNYA 769
QY 677 VTKENTIIINPSENGDTSTNGI 697
DB 770 DIKLDI-----NTGNTYIDGI 785

RESULT 3
I39933
cryptic protein - Bacillus anthracis
C:Species: Bacillus anthracis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999
C:Accession: I39933
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr
A:Reference number: I39933; MUID:89172073; PMID:3148491
A:Accession: I39933
A:Status: preliminary; translated from GB/EMBL/DDSI
A:Molecule type: DNA
A:Residues: 1-192 <RES>
A:Cross-references: GB:M22569; NID:g143280; PIDN:AAA22636.1; PID:g143281

Query Match 6.5%; Score 235.5; DB 2; Length 192;
Best Local Similarity 34.6%; Pred. No. 3e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

T28676
rhoptry protein - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28676; A45521
R:Singha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem. Parasitol. 76, 329-332, 1996
A:Title: Comparison of two members of a multigene family coding for high-molecular mass
A:Reference number: 220507; MUID:97077455; PMID:8920022
A:Accession: T28676
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2401 <SIN>
A:Cross-references: EMBL:U36927; NID:g1041784; PID:g1041785; PIDN:AAB41263.1
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-245, 1990
A:Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co
A:Reference number: A45521; MUID:91101660; PMID:2270106
A:Accession: A45521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 2260-2401 <KEP>
A:Cross-references: GB:M34281

Query Match 5.2%; Score 190.5; DB 2; Length 2401;
Best Local Similarity 21.6%; Pred. No. 0.22;
Matches 176; Conservative 120; Mismatches 321; Indels 197; Gaps 40;

QY 1 EVKQENLLNESSSQGLGYFSDLN-----FCAPVVTSSITGDLIS 45
DB 402 EVAKENVLVNYSNLEIKKYNQINIDNKEKQVNDQFKHMKITPPNEMKYQK 461

QY 46 PSELENIPSENYFQSAL--WSGFIKV---KKSDEYTFATSADNHTVMVDDQEV---I 97
DB 462 PSIEIKIMKDE--FLSKVKNYNDPKVYKVESEHKNFTLTNKKITVESDEBIKKYE 518

QY 98 NYASNSNKIRLEKRLYQIKIYOQRENPEKGLDFKLYWDSQKKVEISSDNQLPLK 157
DB 519 NKFNDKSLINETK--STEEYQNTLTKVDD--YIKVCLNTNELITNCHNQITLK 573

QY 158 QKSSNRK--KRSTAGTPVPDRDNDGIPD-----SLEVEGYTVDVNKKTFPLSP 205
DB 574 DKLNQNIKITEKNSIDKIYTDKFNELTDKKTELETRFTGLSLNHNESNNKELLTYFD 633

QY 206 WTSNIEKGLYKSSPKNSWSTADPSDFKVTGRIDKNVSPARHPLVAAYPIVHVD 265
DB 634 LKANLGNKNMLYKOFNEK---EKAVEDIRKKVNDINKIVSN-----IRIT 677

QY 266 MENILSKNEDOSTNDTSETRTISKTSTSTHTSNSSSTVAI--DHSLSLAGERTWAE 324
DB 678 IYTSIYNINEDTENEIGKSIELLNTKVEKRVANVTNLNEIKELKDYDFQDFGK---- 732

QY 325 TMLGNTADTARLANRINYNTGTAPIYVLPITSVLGNK--OTLTIKAD--ENQLSQI-- 380
DB 733 -----EKNIKYPDEN--KIKNDIDTLNQIKIDSITLTTEIKKNSENHIDEIKG 778

QY 381 -----LAPN-----SKNLAIALNAQ-----DASSTPTMNYQFLE 419
DB 779 QIDKLKVPNKTWFBDDPEIEKKIENIVEKIDKKNIYKEIDKLLNEISKIENDKTSLE 838

QY 420 KTKQLRLDTDOVYGNATYNFENGRVVDTSNWSSEV-----LPOIQTETARIIFNGKDLN 475
DB 839 KLANINLSYKSLGNFLQOIDEKKKAETHIKAMEAYDDLDNKKKSOEI---EKWN 895

QY 476 L-----VERRTAANPSDP-----LETT-----KPDMLKEALKIAGFNEP---NGN 515
DB 896 INMDIKDTHKMKALNISHDDYKIVHTTSKHNHEKISDIRKNSLKIQDFSESYINDI 955

QY 516 LOYQGGDIETFDNFQOQSNTKNQALNATNIYTVLDKIKLAKNMILLRDKRF--H 573
DB 956 KKELEKNVLE-----SQNNNTDINQYLSKIE--NIYNIL--KLNKIKKIIDKVEYTD 1005

QY 574 YDRNIAVGADESVMVKEAHREVINSSTEGLLNIDKDIRKILSYGIVIEDTEGLKE--- 630

Db 1006 IEKNN-----KKINALENS-----EKI-----ITQKENSLEKQCS 1038
QY 631 ----VINDRY---DMNISLRODQKTFIDFKYNDKPLPLYISN--PNYKVN----- 674
Db 1039 KIKSTIDNDYVSECIKNITNL-----KTYIVNEKN--INTYFKABEYQNVSLENNIE 1092
QY 675 YAVTKENTINPSENGDTSN--GIKKILIFSCK 706
Db 1093 MADTKSQYILNKKNGTNTNDYNIKELKHKKK 1126

RESULT 9
G90599
hypothetical protein MYPU 7030 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: G90599
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, P.; Moszer, I.
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul
A:Reference number: A9512; MUID:21267165; PMID:11353084
A:Accession: G90599
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-752 <KUR>
A:Cross-references: GB:AL445566; PID:g14090118; PIDN:CAC13876.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU 7030
A:Genetic code: GCG

Query Match 5.2%; Score 190; DB 2; Length 752;
Best Local Similarity 18.7%; Pred. No. 0.043;
Matches 152; Conservative 141; Mismatches 258; Indels 260; Gaps 41;

QY 41 GDLSISSELENIPSENYFQSALWSGFIKKVKSDEVT-----AASADNHTVMVDD 92
DB 20 GTISGYSIGLSQLPNESASL-----VKADELMPNLKNQSVLNSKISIELINQWQD 69

QY 93 DQEVINKASN-----SNKIRLEKGR--LYQI---KIQYQRENPEKGLDFKLYWDSQ 140
DB 70 ENKVNISANDFFNKVFSQKTPLENGEKITYSVLGRDIFQIWNPSQKITSKVSKITSSK 129

QY 141 NKEVTSNDLQPLPELKQKSSN--SRKSTSTAGTPVPDRDND-----GTPDSLEVEGYTV 194
DB 130 ISKDVWMD-----KQKELNDFAKNLRVNFKGSSASEKQSDIWAASQFNQSKLBIKYL 182

QY 195 DVKN-----KRTFLSPWISNIHEKK--GLTKYKSSPEKMWSTASDPY 233
DB 183 DDKNVEKNISENTEFDELKTEENAFVGGYSNDIAGTKVIEAIVKYKNSEK-----Y 235

QY 234 SDFEKTGRIDKN-----VSPARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRT 288
DB 236 QNIQKI--QITNNFKRFDTSDES-----LRWELSNVKTERTQSKSEN---DASA 280

QY 289 ISKNTSTSRHTSNSSSTVAIDHSLSLAGERTWAEWTGLNTADTARLANIRYVNTGTA 348
DB 281 FSGNSLVSKNSLKNDEKIKLESFV-----FDLKTNENTKLSVKYRDIK---- 327

QY 349 PIYNVLPTTS-----LVLGKNQTLATIKADENQSLQILAPNPNYPSKNLAP 394
DB 328 --FNKLETVGEQKSVKLTLYLIGKVVVDGNEI-----PNRLTSLFRPS---SKESTI 375

QY 395 IALNAQKASSTPTMNYNQFLEKTKQLRLDTDOVYGNATYNEE---NGRVVDTC 450
DB 376 ILTGLKEELSRKLNLSIVDKVELKWTSKDAID--KILASTITKDFQSQSKRITIGPK 434

QY 451 SN-----WSEV--LPOIQTETTA---RIIFNGKDLNMLVERRIAAV--NPSPDPLETTK 494
DB 435 ANTDIAKELAKYLTEISVENINDVGTLYLKVFLKDDSTINRTITLVLGSGFAKVEATK 494

QY 495 PD-----MTLKEALKIATFGFNEPFGNGLQYQGGDIETFDNFQOQSNTKNQALNAT 548

Db 702 NKLSEISKEDNTSLEKVKDINLSY--GQNLGNLFLE---QIDEEKKKAENTIK----- 752
Qy 170 SAGTVPDRDN-----DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPE 224
Db 753 SMEAYIDDLNIIKKKSQEIETEMDIK--MDINKEMEALK--ISHDDDKCHKDKSKNHE 807
Qy 225 KWSTASDP-----YSDFEKVTGRDKDNVSPAEARH-----PLVAAPIYHVVD 265
Db 808 NISDIYDKSKKIIQDFPSREDINDINKLQKRVSESQHNDSINDINOCLNEVANIYNILKN 867
Qy 266 MENIILSKNEDOSTQNTSETRTISKNTSTSRTHTSNSNSTVAIDHLSLAGERTWAET 325
Db 868 KIKKIIDKVKETSE-----IEKNKKNINDELNNSEKVIKIEGDLSLKECKRSKIN- 918
Qy 326 MGLNATADTARLANIRYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQOLSQLAPNN 385
Db 919 -----STLDDKDIDECININ-----VLKKN----- 939
Qy 386 YYPKLNAPIALNAQKASSTPITMNYNOFLEKTKQLRLDQVGYNIATYNPENGRV 445
Db 940 -----ILNEE-----TNIT--NHFKNAEB-----YNKIVLSNFIN-- 967
Qy 446 RYDTGSNNSEVLPOIQTETARIFNGKDLNLVERIAAVNPSDPLETTKPDWTLKALKI 505
Db 968 -LEMADNKSQYILEIKONNGT--NDHDYNIKELK-SHKDKSNGYK-TEADQNKAIQK- 1020
Qy 506 AFGFNEPNGNLOVQKGD-----ITEPDPNDOOTSONIKNOLAEALNATNIYTVLD 555
Db 1021 -----NKELEFYQKEEVTVLLNKYYAVELKNKFD-KTKDKSKOIIEIKDAHNYCTLE 1072
Qy 556 KIKLNAAKN-----ILIPDKRPHYDRNNAIVAGDESIVVKEAHREVN-----SSTEGLLL 605
Db 1073 SGKSEKKNNEIKNEKIHIEDEVANDKSKAITSIKSVPEPKTKIIEIRTKSDDCCL 1132
Qy 606 NIDKDKIRKILSGYVIEIDT-----EGLKEVINDRYDMLNLSLRQDKTFIDPKK 656
Db 1133 KETNDEKQISNLSIDTQETKLTENGKQKLTLEL-----LESKKQKNIEDQKK 1183
Qy 657 YNDKJPLYISN-----PNYKNVAVTKENTIINPSNGDPTSTNGIKK 699
Db 1184 ELDEVNSKIKNIENTVNOHKKNVEIGI--VEKINEIAKTNKQIBSTKELIK 1233

RESULT 12
B89921
hypothetical protein ebhA [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizurani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: B89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <KUP>
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BA842527.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebhA

Query Match 5.1%; Score 187; DB 2; Length 6713;
Best Local Similarity 20.6%; Pred. No. 1.5;
Matches 170; Conservative 127; Mismatches 289; Indels 238; Gaps 41;

Qy 9 LNESESSSQGLGYFSDLNFOAPMVVIS-----STTGDLSIPSPSELENIPSENYFOSA 63
Db 2468 VQAKSDAKANLG-TLTHLNNNAQKDLTSQIBGATTVGVNSVKTQAQDLGAMQRLBSA 2526
Qy 64 IWSGRIKVKKSDEY-----TPATSADNHVMTWDDQEVINKASNNKIRLEKGRLYQIKI 118

Db 2527 I-ANKDQTKASENYIDADPTKTAFDNAIT---QAESYLNKHGHTNKK-----QAVEQ 2576
Qy 119 QYQRENPTEKGL--DFKLYWTDSONKEVISSNLOLPE---LKQSSNSRKRKSTASG 173
Db 2577 AIOQSVISTENALANDANLOCAKTEAQAIDNLTLQTLNTPKLTALKQOVNAQR----- 2628
Qy 174 TVPDRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEK-----KGLTKYKSSPEK 225
Db 2629 -----VSGVT-DLKNATSLLNNAQDLQKQIGDHDHTIVAGGNYTNASPOK 2672
Qy 226 WSTASDPY-----SDFEKVTGRI-----DKNVSPPEARHPLVAAV 259
Db 2673 QGAYTAYNAAKNIVNGSPNVIADVTAAATQVNAASTSLNGDTNLATKQAKDALR 2732
Qy 260 PIYHV-DMENIILSKNEDQSTQNTSETRTISKNTSTSRTHTSNSNSTVAIDHLSLAG 318
Db 2733 QMTHLSDAQQSITGQIDSATQVTVQS---VKDNATNLDNAMNQLRNSIANKDEYKASQ 2789
Qy 319 ERTWAETMGLNTADTARLANIRYNTGTAPIY-----NVLPTTSLVLGKNQTLAT 369
Db 2790 PYVDADTDKQNAVTAITSAE-NIINATSOPTLDPSAVTQAAQVNTNKTALNAQNLAN 2848
Qy 370 IK-----ADENQLSOILAPNNYPSKNLAPIALNAQKASSTPITMNYN-QFLELEKTKOL 424
Db 2849 KKOETTANINRLSHL-----NNAQKQDLNLTQVTAAPNISTVYQVTKAE 2892
Qy 425 RL-----DTDOV-----YGNLAT-----YNFENGVRVDTGS 451
Db 2893 QLDQAMERLINGIQDKQVQKQSVNFTDADPEKGTAVNNAVTAAENIINQANG-----TNA 2947
Qy 452 NWSEVLPIQIE-TTARIIFNGKDLNLVERRI--AAVNPSPDPLET----- 492
Db 2948 NQSOVEAALSTVTTTKQALNG-----DRKVTDAKNANQTLSTLDNLNNAQKGAVTGNI 3001
Qy 493 ----TRPDMTLKALKIATGFPNPNGLQ--YQKDIITEFDNF-----QQTQ 536
Db 3002 NQAHVVAEVT--QAIOAQOELNTAMGNLKNLSLNDKDTLIGSQNFADADPEKKNAYNEAVR 3059
Qy 537 NIKNQLAELNATNIYTVLDKIKLNAAKNILIRDKRPHYDRNNAIVAGDESIVVKEAHREVI 596
Db 3060 NAENILNKSGTNV---PKQVVEAAVN-----QVNTTKAALNGTQNL-EKAKQHA 3105
Qy 597 NSSTEGH--LLNIDDKIRKILSGYVIEIDTEGLKEVINDRYDMLNLS-----LRQ---D 647
Db 3106 NTAIDGLSHLTNAKEALKQLVQOSTTVAEQAQNEQKAN-----NVDAAQDLKQSIAD 3159
Qy 648 GKTDFIDFKXNDKPLVYISNPNYKNVY--AVTKENTIINPSN 689
Db 3160 NATTKQNQNTD-----ASPN-KKDAYNNAVTTAOGIIDQITN 3196

RESULT 13
T31105
hypothetical protein 2 - Haemophilus ducreyi
C:Species: Haemophilus ducreyi
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T31105
R:Ward, C.K.; Lumley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.
J. Bacteriol. 180, 6013-6022, 1998
A:Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
A:Reference number: Z20984; PMID:99030326; PMID:9811662
A:Accession: T31105
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-4919 <NAR>
A:Cross-references: EMBL:AF057696; NID:g3929021; PID:g3929023; PIDN:AA079761.1
C:Genetics:
A:Gene: lspa2

Query Match 5.0%; Score 181; DB 2; Length 4919;
Best Local Similarity 21.3%; Pred. No. 1.9;
Matches 176; Conservative 118; Mismatches 316; Indels 218; Gaps 40;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 7.98439 Seconds
(without alignments)
4636.784 Million cell updates/sec

Title: US-09-848-909A-16

Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKKILFSGKGYEIG 711

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3604	98.8	764	1	PAG_BACAN
2	235.5	6.5	204	1	YP11_BACAN
3	198	5.2	1658	1	YMS7_YEAST
4	173.5	4.8	1830	1	MS11_PLAFK
5	173.5	4.8	1639	1	MS11_PLAFW
6	171.5	4.7	1208	1	PCP1_SCHPO
7	163	4.5	1803	1	YUL3_YEAST
8	162.5	4.5	1024	1	RIP3_MOUSE
9	162.5	4.5	1233	1	YF16_YEAST
10	161	4.4	2867	1	RB22_PLAVB
11	160.5	4.4	1487	1	MD33_YEAST
12	160.5	4.4	2869	1	RB11_PLAVB
13	159.5	4.4	1162	1	EXEN_CLOBU
14	159.5	4.4	1358	1	SIR4_YEAST
15	159.5	4.4	2116	1	MY92_DICDI
16	158	4.3	1790	1	US01_YEAST
17	157	4.3	1000	1	S155_YEAST
18	157	4.3	1029	1	RIF3_RAT
19	156	4.3	1957	1	SPOF_SCHPO
20	155.5	4.3	1167	1	CAGA_HELPJ
21	155	4.2	1385	1	FAT1_SCHPO
22	153.5	4.2	1046	1	SBCC_LACLA
23	152.5	4.2	730	1	GLN3_YEAST
24	152.5	4.2	895	1	RA50_THEVO
25	152.5	4.2	2339	1	RPCL_PLAFW
26	151	4.1	2167	1	BEW2_YEAST
27	150.5	4.1	1005	1	RA50_METJA
28	150.5	4.1	1570	1	P3X1_DICDI
29	150	4.1	3418	1	BR22_HUMAN
30	149.5	4.1	1162	1	EXEN_CLOBU
31	149.5	4.1	2376	1	TAQ3_YEAST
32	149	4.1	719	1	YM41_YEAST
33	149	4.1	976	1	SCPI_HUMAN

ALIGNMENTS

RESULT 1
PAG_BACAN STANDARD; PRT; 764 AA.
ID PAG_BACAN STANDARD; Q9KH69; Q9ROU2;
AC P13423; Q9F5R7; Q9KH69; Q9ROU2;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins-
DE translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].
GN PAGA OR PAG OR PX01-110.
OS Bacillus anthracis.
OC Plasmid pX01.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89172073; PubMed=3148491;
RA Welkos S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppla S.H.,
RA Schmidt J.J.;
RT "Sequence and analysis of the DNA encoding protective antigen of
RT Bacillus anthracis.";
RL Gene 69:287-300(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=28, 33, BA1024, and BA1035;
RX MEDLINE=99214082; PubMed=10197996;
RA Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;
RT "Genetic diversity in the protective antigen gene of Bacillus
RT anthracis.";
RL J. Bacteriol. 181:2358-2362(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=V770-NP1-R / ATCC 14185;
RX MEDLINE=20359347; PubMed=10899854;
RA Cohen S., Mendelson I., Altobum Z., Kobiler D., Elhanany E., Bino T.,
RA Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M.,
RA Kronman C., Velan B., Shaferman A.;
RT "Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus
RT anthracis spore vaccines protect against anthrax.";
RL Infect. Immun. 68:4549-4558(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Steirne;
RX MEDLINE=99445483; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
RA Keim P., Koeher T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
RA Martinez Y., Rieke D., Svensson R., Jackson P.J.;
RT "Sequence and organization of pX01, the large Bacillus anthracis
RT plasmid harboring the Anthrax toxin genes.";
RL J. Bacteriol. 181:6509-6515(1999).
RN [5]
RP DOMAINS.
RX MEDLINE=91332080; PubMed=1651334;
RA Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;
RT "The carboxyl-terminal end of protective antigen is required for
RT receptor binding and anthrax toxin activity.";

Q07833 bacillus su
P41508 mycoplasma
P38537 bacillus sp
P47037 saccharomyc
Q01550 xenopus lae
P54697 dictyosteli
P19321 clostridium
Q92kw5 helicobacte
P08964 saccharomyc
P59368 saccharomyc
P45949 saccharomyc
P25389 saccharomyc

34 149 4.1 2334 1 WAPA_BACSU
35 148.5 4.1 979 1 P115_MYCHR
36 148.5 4.1 1176 1 SLAP_BACSH
37 148.5 4.1 1230 1 SMC3_YEAST
38 148.5 4.1 1744 1 TANA_XENLA
39 148.5 4.1 2245 1 MYSJ_DICDI
40 148 4.1 1276 1 BXD_GLOBO
41 148 4.1 1288 1 VACA_HELPJ
42 148 4.1 1328 1 MYS1_YEAST
43 147.5 4.0 678 1 YNC7_YEAST
44 147.5 4.0 817 1 YG4A_YEAST
45 147 4.0 1037 1 KCC4_YEAST

RL J. Biol. Chem. 266:15493-15497(1991).
RN [6]
RP CHARACTERIZATION.
RC STRAIN=Sterne;
RX MEDLINE=94327640; PubMed=8051159;
RA Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;
RT "Anthrax protective antigen forms oligomers during intoxication of
RT mammalian cells.";
RL J. Biol. Chem. 269:20607-20612(1994).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=21129592; PubMed=11207581;
RA Beaugard K.E., Collier R.J., Swanson J.A.;
RT "Proteolytic activation of receptor-bound anthrax protective antigen
RT on macrophages promotes its internalization.";
RL Cell. Microbiol. 2:251-258(2000).
RN [8]
RP TOXIN REGULATION.
RX STRAIN=Weybridge;
RX MEDLINE=94131936; PubMed=8300513;
RA Koehler T.M., Dai Z., Kaufman-Yarbray M.;
RT "Regulation of the Bacillus anthracis protective antigen gene: CO2 and
RT a trans-acting element activate transcription from one of two
RT promoters.";
RL J. Bacteriol. 176:586-595(1994).
RN [9]
RP MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.
RC STRAIN=Sterne;
RX MEDLINE=95050722; PubMed=7961869;
RA Singh Y., Klimpel K.R., Aroa N., Sharma M., Leppla S.H.;
RT "The chymotrypsin-sensitive site, PFD315, in anthrax toxin protective
RT antigen is required for translocation of lethal factor.";
RL J. Biol. Chem. 269:29039-29046(1994).
RN [10]
RP MUTAGENESIS OF DOMAIN 4 LOOPS.
RC STRAIN=Sterne;
RX MEDLINE=99185012; PubMed=10085028;
RA Varughese M., Teixeira A.V., Liu S., Leppla S.H.;
RT "Identification of a receptor-binding region within domain 4 of the
RT protective antigen component of anthrax toxin.";
RL Infect. Immun. 67:1860-1865(1999).
RN [11]
RP MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.
RC STRAIN=Sterne;
RX MEDLINE=21032804; PubMed=11178978;
RA Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;
RT "Trp 346 and Leu 352 residues in protective antigen are required for
RT the expression of anthrax lethal toxin activity.";
RL Biochem. Biophys. Res. Commun. 281:186-192(2001).
RN [12]
RP MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.
RC STRAIN=Sterne;
RX MEDLINE=21438996; PubMed=11554763;
RA Ahuja N., Kumar P., Bhatnagar R.;
RT "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are
RT required for oligomerization of anthrax protective antigen.";
RL Biochem. Biophys. Res. Commun. 287:542-549(2001).
RN [13]
RP MUTAGENESIS OF PRO-289.
RC STRAIN=Sterne;
RX MEDLINE=21255689; PubMed=11356563;
RA Khanna H., Chopra A.P., Aroa N., Chaudhry A., Singh Y.;
RT "Role of residues constituting the zeta strand of domain II in the
RT biological activity of anthrax protective antigen.";
RL FEMS Microbiol. Lett. 199:27-31(2001).
RN [14]
RP MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.
EX MEDLINE=21125576; PubMed=1122612;
RA McDridge J., Mourez M., Collier R.J.;
RT "Involvement of domain 3 in oligomerization by the protective antigen
RT moiety of anthrax toxin.";
RL J. Bacteriol. 183:2111-2116(2001).
RN [15]

RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.
RX MEDLINE=21269403; PubMed=11113126;
RA Sellman B.R., Nassi S., Collier R.J.;
RT "Point mutations in anthrax protective antigen that block
RT translocation.";
RL J. Biol. Chem. 276:8371-8376(2001).
RN [16]
RP MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;
RX ILE-239; TRP-255 AND PHE-265.
RC STRAIN=Sterne;
RX MEDLINE=22112896; PubMed=12117959;
RA Chauhan V., Bhatnagar R.;
RT "Identification of amino acid residues of anthrax protective antigen
RT involved in binding with lethal factor.";
RL Infect. Immun. 70:4477-4484(2002).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=97192099; PubMed=9039918;
RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;
RT "Crystal structure of the anthrax toxin protective antigen.";
RL Nature 385:833-838(1997).
RN [18]
RP REVIEW.
RX MEDLINE=21428689; PubMed=11544370;
RA Mock M., Fouet A.;
RT "Anthrax.";
RL Annu. Rev. Microbiol. 55:647-671(2001).
CC -1- FUNCTION: One of the three proteins composing the anthrax toxin,
CC the agent which infects many mammalian species and that may cause
CC death. PA binds to a receptor (ATR) in sensitive eukaryotic
CC cells, thereby facilitating the translocation of the enzymatic
CC toxin components, edema factor and lethal factor, across the
CC target cell membrane. PA associated with LF causes death when
CC injected. PA associated with EF produces edema. PA induces
CC immunity to infection with anthrax.
CC -1- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a
CC protective antigen (PA), a lethal factor (LF) and an edema factor
CC (EF). None of these is toxic by itself. PA+LF forms the lethal
CC toxin (LeTx); PA+EF forms the edema toxin (EdTx). PA-63 forms
CC heptamers and this oligomerization is required for LF or EF
CC binding. Once activated, at low pH, the heptamer undergoes
CC conformational changes and converts from prepore to pore inserted
CC in the membrane, forming cation-selective channels.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: The molecule is folded into four functional domains. Each
CC domain is required for a particular step in the toxicity process.
CC Domain 1 contains two calcium ions and the proteolytic activation
CC site. Cleavage of the PA monomer releases the subdomain 1a, which
CC is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is
CC part of the remaining 63-kDa fragment (PA63) and contains the
CC binding sites for LF and EF. Domain 2 is a beta-barrel core
CC containing a large flexible loop that has been implicated in
CC membrane insertion and pore formation. There is a chymotrypsin
CC cleavage site in this loop that is required for toxicity. Domain 3
CC has a hydrophobic patch thought to be a separate domain and shows
CC interactions. Domain 4 appears to be a separate domain and shows
CC limited contact with the other three domains: it would swing out
CC of the way during membrane insertion. It is required for binding
CC to the receptor; the small loop is involved in receptor
CC recognition.
CC -1- PTM: Proteolytic activation by furin or a furin-like protease
CC cleaves the protein in two parts, PA-20 and PA-63; the latter is
CC the mature protein. The cleavage occurs at the cell surface and
CC probably in the serum of infected animals as well; both native and
CC cleaved PA are able to bind to the cell receptor. The release of
CC PA20 from the remaining receptor-bound PA63 exposes the binding
CC site for EF and LF, and promotes oligomerization and
CC internalization of the protein.
CC -1- MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were
CC performed that showed that the residues present in the small loop
CC of domain 4, and not the ones in the large loop, are involved in
CC receptor recognition.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

Query Match 98.8%; Score 3604; DB 1; Length 764;
Best Local Similarity 96.2%; Pred. No. 3.9e-171;
Matches 707; Conservative 1; Mismatches 3; Indels 24; Gaps 1;
QY 1 EVKQENRLNSESSESSQGLGYFDFLNFQAPMVVTSSTGDLSPSELENISENQYF 60
DB 30 EVKQENRLNSESSESSQGLGYFDFLNFQAPMVVTSSTGDLSPSELENISENQYF 89
QY 61 QSAIWSGPIKVKSDPYFATSADNHVTVWVDDDEVINKASNSKIRLEKGRLLQIKIY 120
DB 90 QSAIWSGPIKVKSDPYFATSADNHVTVWVDDDEVINKASNSKIRLEKGRLLQIKIY 149
QY 121 QRENTEKGLDPLKLVWTSQNKKEVIVSSDNLQLELPELKOKSSNRKSTSGAPVPRDN 180
DB 150 QRENTEKGLDPLKLVWTSQNKKEVIVSSDNLQLELPELKOKSSNRKSTSGAPVPRDN 209
QY 181 DGIPDSLEVEGVTVDKNRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240
DB 210 DGIPDSLEVEGVTVDKNRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 269
QY 241 GRIDKNVSPKARHPIVAAYPIVHVDMENIISKNEQSTQNTDSETRTISKNTSTSTHT 300
DB 270 GRIDKNVSPKARHPIVAAYPIVHVDMENIISKNEQSTQNTDSETRTISKNTSTSTHT 329
QY 301 -----SNSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 336
DB 330 SEVHGNAEVHASFDPDGGVSAGPSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 389
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQISQILAPNNYPSKNLAPIA 396
DB 390 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQISQILAPNNYPSKNLAPIA 449
QY 397 LNAQKDAASPTITMNYNQFLELEKTKQLRLDQVYGNATYFNENGRVVDGTGNSWSEV 456
DB 450 LNAQKDAASPTITMNYNQFLELEKTKQLRLDQVYGNATYFNENGRVVDGTGNSWSEV 509
QY 457 LPQIQTETARIIFNGKDLNLVRRRIAANVPSDPLETTPDMLKEALKIAFGFNEPKNL 516
DB 510 LPQIQTETARIIFNGKDLNLVRRRIAANVPSDPLETTPDMLKEALKIAFGFNEPKNL 569
QY 517 QYQKDIITEFDNFQOQTSQNKQLAELNATYVLDKIKLAKKMLIRDKRFHYDR 576
DB 570 QYQKDIITEFDNFQOQTSQNKQLAELNATYVLDKIKLAKKMLIRDKRFHYDR 629
QY 577 NNIAVGADESIVKAEHREVINSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 636
DB 630 NNIAVGADESIVKAEHREVINSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 689
QY 637 DMLNTSSLRQDGTFTDFKKNYDKLPLYSNENYKVNVTAKNTIINPSENGDTSTNG 696
DB 690 DMLNTSSLRQDGTFTDFKKNYDKLPLYSNENYKVNVTAKNTIINPSENGDTSTNG 749
QY 697 IKKILIFSKKGYEIG 711
DB 750 IKKILIFSKKGYEIG 764
RESULT 2
YPB1_BACAN
ID YPB1_BACAN STANDARD; PRT; 204 AA.
AC P13422; Q9X377;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein pXOI-111.
GN pXOI-111.
OS Bacillus anthracis.
OC Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.

MEDLINE=89172073; PubMed=3148491;
Wolkos S.L., Lowe J.R., Eden-McCurchan F., Vodkin M., Leppia S.H.,
Schmidt J.J.;
RT "Sequence and analysis of the DNA encoding protective antigen of
Bacillus anthracis.";
RL Gene 69:287-300(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Storne;
RC MEDLINE=99445483; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
RA Kaim P., Koeher T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
RA Martinez Y., Riecke D., Svensson R., Jackson P.J.;
RT "Sequence and organization of pXOI, the large Bacillus anthracis
plasmid harboring the Anthrax toxin genes.";
RL J. Bacteriol. 181:6509-6515(1999).
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22589; AAA22636.1; -;
DR EMBL; AF065404; AAD32415.1; -;
DR PIR; G59104; G59104.
DR PIR; I39933; I39933.
DR HSP; P33423; IACC.
KW Hypothetical protein; Plasmid; Transmembrane.
FT TRANSMEM 162 182 POTENTIAL.
FT CONFLICT 93 93 T -> A (IN REF. 1).
FT CONFLICT 184 204 KSCNCLIIYVEVSQMLNSVFY -> NHVIVILSM (IN
FT REF. 1).
SQ SEQUENCE 204 AA; 23029 MW; E1657B23AB4273FD CRC64;
Query Match 6.5%; Score 235.5; DB 1; Length 204;
Best Local Similarity 34.6%; Pred. No. 2.2e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;
QY 563 MNTILDRKPHYDRNNIAVGADSVVKEAHREVINSTEGLLNIDKIRKILSGYVIEI 622
DB 1 MNTILDRP-YHYDNGNIGVDSDYLKNAKQILWSSDGVSLNLDENVQALSQYMLQI 59
QY 623 EDTE-----GLKEVINDRYDMLNTSSLRQDGTFTDFKKNYDKLPLYSNEN 669
DB 60 KKP-SNHLTNSPVTITLAGKDSGVGELYRVLS-----DGTGFLDNKFDENWRSLV-DPG 112
QY 670 YKVNVTAKNTIINPSENGDTSTNGIKKILIFSKKGYEIG 710
DB 113 DDVYVAVTKEDFNAVTRDENGNI-NKLNKTLVLGKIKEI 153
RESULT 3
YM67_YEAST
ID YM67_YEAST STANDARD; PRT; 1658 AA.
AC Q03661; Q04998;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 187.1 kDa protein in GUA1-ERG8 intergenic region.
GN YMR219W OR YMR261.13 OR YMG959.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=9733268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,

RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moulle S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RL Nature 387:90-93(1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z49809; CA489934.1; -;
DR EMBL; Z49939; CA90190.1; -;
DR PIR; S55101; S55101.
DR Germonline; 142894; -;
DR SGD; S0004832; ESC1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
KW Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;

Query Match 5.2%; Score 188; DB 1; Length 1658;
Best Local Similarity 19.1%; Pred. No. 0.08;
Matches 159; Conservative 139; Mismatches 275; Indels 260; Gaps 40;

QY 1 EVKQENLLNESES---SQGLGYFSDLAFAQPMVTSITGDLSPSELENIPISEN 57
DB 290 EMELEDDIDVEDAKESQGAEGTEHS-VDFSKYMQPRTDNTKIPVIEKYESDEHKVHQ 348
QY 58 QYFQSAINS-GPIKVKKSD-----YTFATSAADN---HVTMWVDDQEVINKASNSNK 105
DB 349 RYSEGDAPFGSVNIVSDDESEDESAESYSANAENVYHNEHLEDDKELIEDIESDS 408
QY 106 ILEKGRLYQIKIQRENPEKGLDFK-----YWTDSQNKKEV 145
DB 409 -----ESQASESQGSDDFEYKVKKEKSTSEETENTSESRDQGFADYATNKV 459
QY 146 ISSDNLQPEL-----KQSSNRKRRKSTSGAPTPVRDNDGIPDSLEVE 190
DB 460 EQQENDEPERKDIIRSLDKNFHGNKSEYSENVLENETDPAIVERNO-IND---VE 515
QY 191 GYTVDVQKRTPL-----SPWISNIHE--KKGLTKYKS-----PEKWTASD- 231
DB 516 GYDVTGKSVESDLHSP--DNLYDLAARAMLQFOOSRNSNCPOKEQVSESYLHGSNGS 573
QY 232 -----PYSDFEKTGTGRIDKNVSPFAHPLVAAVPIVHVDMENIILSKNEDQ 277
DB 574 NLSGRSLDESEEQIPLKDF---TGNNNNLKTID-RGDLSS--VELEVEKY---SEKK 622
QY 278 STQNTDSETRTSKNTSTSRHTSNSSTVAIDHSLSLAGERTWAETMGANTADTARLN 337
DB 623 LDGSTKEKELVPLSTDTTINNSLGNEDSIYYSLOD-----ADAISENLTD----- 667
QY 338 AMRVYVNTCTAPVYNVPLTSLVLCKNQTLATIKADENCLS-----QILAP----- 383
DB 668 ---VPLMEKTKTKYEVVISEV-----YSTSYEDNTVAMPQVEYTSFPMNDPNSL 718
QY 384 NNYPSKN-----LAPIALNAQKQAS-----STPTIMYNOFLEKTKQL 424
DB 719 NDDYEKKHDLKSLTALAAPAF--TKDAEFVEAGVTKSLTSTSGHTNIPTSSETKQV 776
QY 425 RLDTDOVGNATYNFENGRVVDVGSNWSVLPOIQETTARIIFNGKDLNLVETAAV 484
DB 777 S-DLDESTENVTFENGTG-----DENKQSKNFPQVANSTDKSTEDNTD-----EKYFSAI 827
QY 485 NPSDFLETKPDWTLKALKAFGNENPGLQYQCKDITBDFNPF-DQQTSON---IKN 540
DB 828 NVTN-----VTGSSCEDIIETASNVEE---NURYCEKDMNEARMSGDCYKQNDGSKT 880

QY 541 QLA-----ELNATNIYTVLDKIKNAKQNLIRDKRPHYDRNNIAGADESVVKEAHR 593
DB 881 QISFSTSDPNFOESNDNTEFSSTK-----YKVRNSDLDEDSLKKELTAK 925
QY 594 -EVIN-----SSTREGLLNIDKDIRKILSGVIVIEITEGKEVI 632
DB 926 AEVVDKLDRESEDSVEQDYADPEPQNDGSENVKGTKK-----DTLGIVEPE 975
QY 633 NDRYDMLNLSURQDQKTFIDPKYNDKLPYISNPNKYNYAVYTKENTIIN 685
DB 976 NEXVN-----KVHEETLFEANVSSSVVQNKDMHTDVIN 1010

RESULT 4
MSPI_PLAFK STANDARD; PRT; 1630 AA.
ID MSPI_PLAFK AC P04932;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merzoite surface protein 1 precursor (Merzoite surface antigens)
DE (PMMSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
RA Stunnenberg H., Bujard H.;
RT "Polymorphism of the precursor for the major surface antigens of
RT Plasmodium falciparum merzoites: studies at the genetic level.";
RL EMBO J. 4:3823-3829(1985).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RA Pan W., Rolie K., Bujard H.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- PTM: Merzoite surface antigen contain the sequence of 83 kDa, 42
CC kDa and 19 kDa antigens which are the major surface antigens of
CC merzoites. The maturation take place during schizont.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X03371; CAA27070.1; -;
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merzoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).

```
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;

Query Match
Best Local Similarity 4.8%; Score 173.5; DB 1; Length 1630;
Matches 168; Conservative 134; Mismatches 251; Indels 263; Gaps 47;

QY 5 ENRL-LNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQVFOQA 63
Db 918 ENILSGKNKIYQELIGQKSE-NF-----YKILKXSDTFYNE 956
QY 64 IWSGFIKVKSDYFATSAADNHVTWDDQEVINKASNS-----NKIRLEKGRLY- 114
Db 957 SFTNFVKSADD-----INSLNDESKRKKLEEDINKLKTQLQSPDLNFKYKLERLFD 1011
QY 115 -----QIKIOYORENPTKGLDFKLYWTDQNKKEVISSDNLQLPKQKSSNRKKS 168
Db 1012 KKTIVGKYKQIKLTLKEQLESKL--NSLNPKHVL--QNFVFFNKKAEAEIATEN 1076
QY 169 TSAGTPVPRDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGILTKY---KSSP-- 223
Db 1068 T-----LENYKILLKH-----XGLVKYNGESSPLK 1094
QY 224 ---EKWSTASDPYSDFE---KVTGRIDKNVSPAR-----HPLVAAYPIVHVD 265
Db 1095 TLSESIQTDNYASLENFKVLSKLEGLKONLLEKKLSYLSGLHLLIA-----E 1147
QY 266 MENILSKNEDOSTQDSTRTISKTSTST-----RHSTNSNSTVAIDHSL 314
Db 1148 LKEIVKYN---YTGNSFSENVDVNALESVKFLPREGTVDATVVSSEGGDTLEQSQPK 1204
QY 315 SIAGERTWAETGLTADTAPLANIRYNTGTAPIYVNLPTT-----SLVLGRNQTLAT 369
Db 1205 KPASTHVGAE---NTITTSQ-NVDEVDVVIIPFGESEEDYDDLQGVVTEAVTSPV 1260
QY 370 IKADENQLSOLLAPNYPSPKULAFIA---LNAQKDASTITWYN-----QFLE 419
Db 1261 I---DNILSKI---ENEVEVLKPLAGYSLKQLENNVTNFKVNDILNSRFNKRE 1314
QY 420 KTKQLRLDTPQV-YGNIAI-----YNPENGVRVDTGNSWSEVLFPQIQTARIIFN 470
Db 1315 NFKV-LESDDLPIYKDLTSSNVVVVDPKYKFNKEKRDKFLSSYVVKDSID----- 1364
QY 471 GKDNLVVERIAAVNPSPL-----ETTKPDM-TLKEALKIAGFNFPN-----GNLQY 518
Db 1365 -TDINFA-----NDVLGYKILSEKYSKSDLSISIKYINDKQGENEKYLPFLNNIET 1414
QY 519 QGKDI-----FDNFDDQTSQNIKNQALNATNIYVLDIKLNAXKN- 564
Db 1415 LVKTVNDKIDLVFVHLEAKVLNYTEK---SNVEVKIKELN--YKLTQDKLADFKNKN 1469
QY 565 -ILIRDKPHYDRN-----IAGV-ADSVKVEAHRVINSSTEGLLINIDK--IRKIL 615
Db 1470 FVGIADLSTDYNNHNLTKFLSTGMVFENLAKTVLSNLIDGNLQ--MLNISQHCQVK-- 1526
QY 616 SGYIVIEIDTEGLKEVINDRYDMLNLSLRDQGTFFIDFKYNDKLPYISNPNKYNVY 675
Db 1527 -----QCQNSGCFPHLDE-----RECKCLNFKQEGKC---VENPAPTGN-- 1566
QY 676 AVTKENTINPENG-----DSTNGIKKI 700
Db 1567 -----EN-----NGGCDADAKCTBEDSGSNG-KKI 1590

RESULT 5
MSPL_PLAFW
ID MSPL_PLAFW STANDARD; PRT; 1639 AA.
AC P04933;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMSEA) (P195).
```

```
GN MSP-1.
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86014355; PubMed=2995820;
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
RA Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
RA Freeman R.R.;
RT "Primary structure of the precursor to the three major surface
RT antigens of Plasmodium falciparum merozoites."
RL Nature 317:270-273(1985).
RN [2]
RP REVISIONS.
RA Holder A.A.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -2- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42
CC kDa and 19 kDa antigens which are the major surface antigens of
CC merozoites. The maturation take place during schizont.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X02919; CAA36676.1; -.
DR PIR; A24594; A24594.
DR PIR; S05603; S05603.
DR PDB; 1CEJ; 28-MAY-99.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1639 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 116 116 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 768 768 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 783 783 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 844 844 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 920 920 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;

Query Match
Best Local Similarity 4.8%; Score 173.5; DB 1; Length 1639;
Matches 168; Conservative 134; Mismatches 251; Indels 263; Gaps 47;

QY 5 ENRL-LNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQVFOQA 63
Db 927 ENILSGKNKIYQELIGQKSE-NF-----YKILKXSDTFYNE 965
QY 64 IWSGFIKVKSDYFATSAADNHVTWDDQEVINKASNS-----NKIRLEKGRLY- 114
Db 966 SFTNFVKSADD-----INSLNDESKRKKLEEDINKLKTQLQSPDLNFKYKLERLFD 1020
QY 115 -----QIKIOYORENPTKGLDFKLYWTDQNKKEVISSDNLQLPKQKSSNRKKS 168
Db 1021 KKTIVGKYKQIKLTLKEQLESKL--NSLNPKHVL--QNFVFFNKKAEAEIATEN 1076
QY 169 TSAGTPVPRDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGILTKY---KSSP-- 223
```

Db 1077 T-----LNTKLLKHV-----KGLVYKNGESSPLK 1103
Qy 224 ---EKWSTASDPYSDPE-----KVTGRIDKNVSPPEAR-----HPLVAAPVIVHD 265
Db 1104 TLSESIOTEDYASLENPKVLSKLEGLKONLNLEKKKLSYLSGLHLLIA-----E 1156
Qy 266 MENIILSKNEDQSTQNTSETTISKTNTS-----RHTSNSTSTVAIDHSL 314
Db 1157 LKXVKNKN---YTGNSPENNNTDYNNALESYKXKLPBGTDVATVVSSESDTLEQSQPK 1213
Qy 315 SIAGERTWAETMGLNTADTARLANAIRVYNTOTAPIYVLPPT-----SLVLGKNTLIAT 369
Db 1214 KPASTHVGAES---NTITTSQ-NVDEVDVDDVLIIVIFGESEEDYDLDQGVVTGEATPSV 1269
Qy 370 IKADENQLSQILAPNYPSPKSLAPIA---LNAQXDAASSTPTIMYN-----QFLELE 419
Db 1270 I---DNILSKI---ENEYEVLYKPLAGVYRSIKKQLENNVMTFVNVKDIILNSFEKRE 1323
Qy 420 KYQRLRLDQV-YGNIAT-----YNFENGRVVDVTGNSWSEVLPOIQTETARIIFN 470
Db 1324 NFKNV-LESDLIPYKDLTSSNVYKDPYKFLNKEKDKFLSSYNIKDSID----- 1373
Qy 471 GQDLNLVERRIAAVNPSPDL-----ETTKPDM-TLKEALKIARFGENEPN-----GNLQY 518
Db 1374 -TDINFA-----NDVLGVYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLNNIET 1423
Qy 519 QGKOITE-----PQNFQDQTSQNIKNQLAPLNATNTYVLDKIKLNAGVN- 564
Db 1424 LYKTVNDKIDLFVIHLEAKVLNVTYVEK---SNVEVKIKELN---YLKTIQDKLADPKKNN 1478
Qy 565 -ILIRKPRHYDRNN---IAVG-ADESUVKEAREHVNSTEGILLNIDKO---IRKIL 615
Db 1479 FVGIADLSTYHNHLLTFLSGVFNELAKTV-SNLDGMLQG-MNISHQCVKXK-- 1535
Qy 616 SGVIVEIEPTGLKEVINDRYDMNLISLRQDKTFIDPKYNDKPLYSINPNYKNVY 675
Db 1536 ---QCPQNSQCFRHLDE-----RECKCLLNKQEGDKC---VENPNPTCN-- 1575
Qy 676 AVTKENTINPSNG-----DTSNGIKKI 700
Db 1576 ---EN-----NGCDDAKCTEEDSGSNG-KKI 1599
PCPI1 SCHPO STANDARD; PRT; 1208 AA.
AC Q2351; 1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Spindle pole body protein pcpl.
GN PCP1 OR SPAC6G9.06C.
CS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=21852775; PubMed=11864908;
RA Flory M.R., Morpew M., Joseph J.D., Means A.R., Davis T.N.;
RT "pcpl, a Spcl10p-related calmodulin target at the centrosome of the
RT fission yeast Schizosaccharomyces pombe";
RL Cell Growth Differ. 13:47-58 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
Waltjens I., Vansreels E., Rieger M., Schaefer M., Mueller-Auer S.,
Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
Borzym K., Langer I., Beck A., Leirach H., Wambutt R., Purnelle B.,
Rager P., Zimmermann W., Wedler H., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
Domiguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
Carutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880 (2002).
CC -!- FUNCTION: Spindle pole body component that binds calmodulin.
CC Overexpression of pcpl causes the formation of supernumerary SPB-
CC like structures and disrupts both mitotic spindle assembly and
CC chromosome segregation.
CC -!- SUBCELLULAR LOCATION: Spindle pole body.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL; Z81317; CAB03608.1; -;
DR EMBL; AF348506; AAX31344.1; -;
DR GenBank; T39068; T39068.
DR GeneDB; SPombe; SPAC6G9.06C; -;
KW Calmodulin-binding; Coiled coil.
FT DOMAIN 151 375 COILED COIL (POTENTIAL).
FT DOMAIN 387 803 COILED COIL (POTENTIAL).
FT DOMAIN 874 1091 COILED COIL (POTENTIAL).
FT DOMAIN 1177 1204 COILED COIL (POTENTIAL).
SQ SEQUENCE 1208 AA; 140763 MW; 70264159ADD42424 CRC64;
Query Match 4.7%; Score 171.5; DB 1; Length 1208;
Best Local Similarity 20.0%; Pred. No. 0.34;
Matches 151; Conservative 132; Mismatches 328; Indels 145; Gaps 31;
Qy 10 NESSESSQGLLG-----YYFSDINFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 78 NKNSDKYNGSLGDKGSPDPSYGLSALKQATQCEALISQGNDSYDVSKLTDLKNSKI- 136
Qy 61 QSAIWSGFIKKVKSDEYTFATSNADNHVMTVMVDQEVINKASNSN-----KI-----RLEK 110
Db 137 -----DHTDGLPANAALTIREQEVLEKVSRENFGRLIKIVCLEKRLS 181
Qy 111 GRLVQIK-----IQYQRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSN 162
Db 182 MAEQIKVAVKDNVELHAERAN-----LQIQKRTES-----LLQKSED 220
Qy 163 SRKRSTRSAGFTVPRDNDGIPDSLEVEGYTVDKVKRTFLSPWISNIHEKGLTKYKSS 222
Db 221 KNFKLEKV--DYLSKVDN-VEQSQNVKVFETIR-----FLENALKVQREKDSLSTEME 273
Qy 223 PEKWTASDPSYDFEYKVTGRIDKNVSPFAHPLVAAPVIVHVDMENTILSKNEDQSTQNT 282
Db 274 EDKSNKEVDVEYERIQNLRLDE--LSE-----LDVAQDLTEKDETA--- 316
Qy 283 DSETRTISKTSTSRHTSNTSNTSSTVAI--DHLSLAGERTWAETMGLNTADTARLAN- 339

Db 317 -TLKROIEKENSSAFENBESSVYVHLOEDYAILQAKCDEFADRIQVLTADLEKENQ 375
QY 340 IRYVNTGTAPIYVLPPTSLVGNKQTLATIKADENQSLQILAPNNYPSKNLAPIALNA 399
Db 376 IMHSEASIGTDSMOVHTLOEQHKAEEIEFLDQISRMNEEGKNFEDIMLOFASLEE 435
QY 400 QKDASSPTITWYVNOFLEKTKRLDQDOVGNATYVNFENGRVVRVDTGNSWSEVLPQ 459
Db 436 ERDVLESKL-----QTLE-DNNSRLMTWSLGNQIESLQNR--EIDBEKHRLLAS 487
QY 460 IOETTARIIFNGKOLNVERIAAVNPSPLETTKPDWTLKBAKIAFGNENPNQYQ 519
Db 488 -----KNSDKALAEATNRLQEVTKLETLR--MKNNDLNEIHDRENEGLTK 535
QY 520 GKDITEFDFNDQOTSONIKH--QLAELNATNITVTLDKIK-----LNAMNIL-LRD- 569
Db 536 IDSITKEKORLINELEQRIKSYEVNSELNGT-IDEYRNKLKXKEEYNEVMAFOYKDN 594
QY 570 --KRFPHYDRNNAVGADESVVYKAEHREVINSTSEGLLNIDKO---IRKILSGYVIEIED 624
Db 595 DLRRFHESINKLOREKELTSLNLEKKNLVISLRETVMLEKERESIKKYLGNKADLDN 654
QY 625 TEGLEKEVINDRY-----DMLNITSLRQDGKTFIDPKYNDKLPYVINSNPYKVN 673
Db 655 T-NLMEILNDKISVLQRLTDVKDELQVSEERE-EAIVAGQKUSASFEL-MSNEKQALE 711
QY 674 VYAVTKENTIINPSNGSTSTNGIKKILIFSKGYE 709
Db 712 LKYSKLNELINAQNLDRREELSEL---SKLFE 744

RESULT 7
YJL3 YEAST
ID YJL3 YEAST STANDARD; PRT; 1803 AA.
AC P47024; P87192;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transposon Ty4 207.7 kDa hypothetical protein.
GN Ty4 OR YJL13W OR J0780.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=97103775; PubMed=8948101;
RA Cziopluch C., Kordes E., Pujol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,
RT SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CTF7, GZF3, two tRNA genes,
RT three remnant delta elements and a Ty4 transposon.";
RL Yeast 12:1471-1474(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49389; CAA89409.1; --
DR SGD; S0003649; YJL113W.
DR InterPro; IPR001584; Rve
DR InterPro; IPR001878; Znf_CHC.
DR Pfam; PF00665; rve; 1.
DR SMART; SM00343; Znf_C2HC; 1.
KW Transposable element; Hypothetical protein.
SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;

Query Match 4.5%; Score 163; DB 1; Length 1803;

Best Local Similarity 19.6%; Pred. No. 1.5;
Matches 161; Conservative 118; Mismatches 262; Indels 282; Gaps 41;
QY 1 EVKQENRLINSESSSQGLGHYFSDNLNFOAPMVVSTTGDLSIPSELENIPSENYF 60
Db 1078 QLKTN--HETSPFKESIG--TNVFRNTNNEISLKTGDTSPITLESINHH-- 1128
QY 61 QSAIWSGFIVKVKSDYTFATSNADNHTVMWDDQEVINKAGNSNKIRLEKGRLYQIKY 120
Db 1129 -----SDY-----STNKVE-----KF 1140
QY 121 QREN-----PTEKGLDFKLYWTDSQNKKEVISDNQLPDLKOKSSNRKSTSGPTVP 176
Db 1141 EXENHHPPIEDIVD-----SDQTDMSNCQDGNLKLKLVTDKNVPTDNGTNSPR-- 1193
QY 177 DRDNDGIDPSLVEGVTVDVKNKRTFLSPWISNIHEKK-----GLTKY----- 219
Db 1194 -----LEONIASPSPQTVNKSFLNKEFSSLNKKEKHEKDNKNSITSVELEKDKR 1247
QY 220 -----KSPKWKSTASDPYSDFEKTGRIDIKQNSPEARHPLVAAYPIVHVDMENILSK 273
Db 1248 SKNRVKLIPDNMTVSAPKIRAIYVNEAISNPDLEKHEYKQAY--HKELQNLKDKM 1304
QY 274 NEDQSTQNTDSE-----TRT-----ISKNTSTSRHTSNS--STVAIDHS 313
Db 1305 VFDVDVKYSRSEIPDNLVPTNTITFKKNGIYKARIVCRGDTQSPDTYSVITTESLNH 1364
QY 314 -----LSLAGERT-WAETMGLNTADT-ARLNANIRYVNTGTAPIYVNLVLP-TTSLVLGNQ 365
Db 1365 HKIFELMIANNRNMFMKLDINHAFYAKLEBE-----IYIPHPRRCVYKLNK 1414
QY 366 TLATIKADENQ-----LSOI-LAPNNYV-----SKNLAPALNAQKSDSTPTM 410
Db 1415 ALYGLKQSPKENDHLRQVNLGIGLKDNTYTPGLVQTEDKNLM-IAYVDDCVIAASNEQ 1473
QY 411 NYNQEL-ELEKTKQLR-----LDTDOVYGNATYNFENGRV-----RVDTGS 451
Db 1474 RLDEFINKLSNFELKITGTLDDVLDTD-ILGMDLVYKRLGTDLTLKSPINRMD--K 1530
QY 452 NMSEVLPOQTETARIIFNGKOLNVERIAAVNP-SDPLETTKPD-----MTLKEALKI 505
Db 1531 KYNEELKIRKSIHPHMYTK-----IDPKQDVLMSEBEFFRGVLKLOQLL-- 1577
QY 506 AFGNEPNNGNLOVQGDITEFDFNDQOTSONIKKQLAELNATNITVTLDKLNAKMI 565
Db 1578 -----GELNVV-RHKCRYDIEFAVKVAVLVNYPHERVFMVYKIIQYL-----V 1621
QY 566 LIRDKRFHYDRN-----NIAVGADESIVKAEHRE-----VI-----NSSTEGLL 604
Db 1622 RYKDIGIHYDRDCNKDKKVIATDASVGSSEYDAQSRIGVILWYGMNIFNVYGNKSTNRCV 1681
QY 605 LNTDKDIRKILSGY-----IVEIEDT-----EGLK----- 629
Db 1682 SSTEAEHALYEGYADSEFLKVLKELGSDNNDIVMTDSKPAIQGLNRSYQPKKEFT 1741
QY 630 -----EVINDRYDMLNTSSLRQDGKTFI-----DFKKY 657
Db 1742 WIKTEIIEKIKESIKLLKITGKNIAIDLLATKPVASDFKRF 1784

RESULT 8
RIP3 MOUSE
ID RIP3 MOUSE STANDARD; PRT; 1024 AA.
AC P97434;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Rho-interacting protein 3 (p116RIP) (RIP3).
GN RHOIP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=97344280; PubMed=9199174;
RA Gebbink M.F.B.G., Kranenburg O., Poland M., van Horck F.P.G.,
RA Houssa B., Moolenaar W.H.;
RT Identification of a novel, putative Rho-specific GDP/GTP exchange
RT factor and a Rho-binding protein: control of neuronal morphology.;
RL J. Cell Biol. 137:1603-1613(1997).
CC FUNCTION: Rho-binding protein involved in control of the actin
CC cytoskeleton. Overexpression promotes neuronal cell flattening and
CC neurite outgrowth, probably by counteracting rhoA-mediated
CC signaling.
CC TISSUE SPECIFICITY: Highly enriched in the brain.
CC SIMILARITY: Contains 2 PH domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; U73200; AAB18198.1;
FIR; T30869; T30868.
MGI; MGI:1349438; Rhoip3.
InterPro; IPR001849; PH.
Pfam; PF00169; PH; 2.
SMART; SM00233; PH; 2.
PROSITE; PS50003; PH DOMAIN; 2.
Guanine-nucleotide releasing factor; Repeat; Coiled coil.
FT DOMAIN 43 150
FT DOMAIN 386 482
FT DOMAIN 164 172
FT DOMAIN 185 190
FT DOMAIN 285 296
FT DOMAIN 671 976
FT DOMAIN 1024 AA; 116363 MW; A918DBB955245E02 CRC64;
SQ SEQUENCE 1024 AA; 116363 MW; A918DBB955245E02 CRC64;
Query Match 4.5%; Score 162.5; DB 1; Length 1024;
Best Local Similarity 19.2%; Pred. No. 0.77;
Matches 157; Conservative 131; Mismatches 304; Indels 225; Gaps 39;
QY 31 APWVTSSTG---DLSPSSSELENIPSENYFOSAIWSGFIKVKSDVTFATSDNH 86
173 AKMAVTSSTGSGSSSSIPSA-EKVPTT---KSTLQEMRAKDPDGTSLSPAQSP 226
87 VTMM-----VDDQEV---NKASNSKIRLEK--RLYQIKIYORENPTEK 128
227 SQSQPPAACTPREPGLESKEDESTISGDRVDGGRKVRVESGYFSLEKAKQLRAEQLPP 286
129 GLDFKLYWTDSONKKEVISS-DNLQLPE-----LKQKSSNSRKKRS----- 168
287 LLSPPSPSTPHRRSQVIEKFALDIEKAHEMETNMLITPSSDTQGRERRAIPKR 346
169 --TSAGTVPDRNDG:PDLSLEVEGYVDVKNKRTFLSPWISNIHEKKG-LTK-YKSSPE 224
347 DPASEAPTAP--LSDACPLSPHRRAKSLDRRTSSMTPLDNF--KKGWLTQYEDGQW 402
225 K--WSTASDP---YSD-FEKTGRIDXNVSPSEARHPLVAAYPIVHVDMENILSKNEDQ 277
403 KKHFWLADQSRUYRDSVAEAAADLGEINLSTCYD-VTEYPV-----ORNYGF 451
278 STONTSETRTSKNTSTSRTHTSNSSTVAIDHSLSLAGERTWAETMGLNTADTARLN 337
452 QIHTKEGEF-TLSAMTSGIR-----RNMIQII----- 477
338 ANIRYNTGTAPYVLPPTSLVLGNQTLATIKADENQLSAILAPNNYPSKNLAPAL 397
478 --MKHVLPAAPDV---TSSLPEGNKGTSPETCSRSTKEQAEPPGPDPEQKKSARE 531
398 NAQKDAASST-----PIWMYNNQ-----FLELEKTKQ----- 423

532 RRREGSKTFDWAEPFPIQALQAQRASAVGSSDSGPGCLEAEPGELERRARRRBP 591
424 -----LRLDTDQVYGNIAFYNGRVRVDTGNNWSEVLPIQIETT- 464
592 KRFGMLDTIDGPMEDTALRMDIDRSPLGLGPPDKTQNVHVEIQRWQV-----ETTP 646
465 -----ARIIFNGKD-----LNLVERRI--AAVNPSPLETTKPDMTLKEAL 503
647 LREKQVPIAPLHLSLEDRSERLSLTHLTSLEKEQSQKEASDLLEQNR---LLQDL 703
504 KIAFGNENGLQYQKDXITFDNFQOOTSQNIKNQLAELNATNITVLDKIKLNAM 563
704 RVALG-REQSAREGYVLQATCRGFAAMEETHQKIEDLQROHQRELE-----KREK 756
564 NILIRDRPHRYDRNNIAVGADESVVKEARE-----VINSSTEGLLNIDKDIRKILSGYI 619
757 DLLAAEE-----TAATISAIEA-MKNAREHREWERELEKSQRSQISSINSIDIALRQYL 909
620 VEIETDEGKEVINDRYDMLNT-----SSLQDQKTFIDFKYNDKPLIYISNPNKYN 673
810 EELQSVQRELEVLSEQYSQKLENAHLAQALRAEQALRCQORENQELNAHQELNNRL- 868
674 VVAVTKENTIIINPSENGDTSTNGIKKILIFSKGYEI 710
869 AARITRLTLTLDGGGEST--GLP--LTQGKDAYEL 901
RESULT 9
ID YF16 YEAST STANDARD; PRT; 1233 AA.
AC P43537;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 137.7 kDa protein in UGSI-FAB1 intergenic region.
GN YFR016C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.;
RL Nat. Genet. 10:261-268(1995).
CC SIMILARITY: SOME, TO MAMMALIAN NEUROFILAMENT TRIPLET M PROTEIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; D50617; BAA09255.1; --
FIR; S56271; S56271.
MGI; MGI:140170; --
DR SGD; S0001912; YFR016C.
DR Hypothetical protein.
KW SEQUENCE 1233 AA; 137697 MW; C8A7CD2C6F0892F6 CRC64;
Query Match 4.5%; Score 162.5; DB 1; Length 1233;
Best Local Similarity 18.9%; Pred. No. 0.98;
Matches 151; Conservative 111; Mismatches 253; Indels 285; Gaps 34;
QY 40 TGDLSI---PSELENIPSENYFQSAIWSGFIKVKSDVTFATSDNHVTMWDDQEV 96

RT structural features with PVRBP-1 and the Plasmodium yoelii 235 kDa
RT thophry protein family.",
RL Mol. Biochem. Parasitol. 108:257-262(2000).
[2]
RN SEQUENCE OF 1189-2439 FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites".
RL Cell 69:1213-1226(1992).
CC -!- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to
human reticulocyte cells.
CC -!- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF184623; AAF7525.1; -.
DR HSP; P03069; IGM.
KW Malaria; Receptor; Signal; Transmembrane; Repeat.
FT SIGNAL 1 21
FT CHAIN 22 2867 RETICULOCYTE BINDING PROTEIN 2.
FT DOMAIN 22 2805 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2805 2826 POTENTIAL.
FT DOMAIN 2827 2867 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 44 133 ASN-RICH.
FT DOMAIN 560 758 LYS-RICH.
FT DOMAIN 1112 1285 LYS-RICH.
FT DOMAIN 2758 2785 7 X 4 AA TANDEM REPEATS OF H-D-D-T.
FT REPEAT 2758 2761 1.
FT REPEAT 2762 2765 2.
FT REPEAT 2766 2769 3.
FT REPEAT 2770 2773 4.
FT REPEAT 2774 2777 5.
FT REPEAT 2778 2781 6.
FT REPEAT 2782 2785 7.
SQ SEQUENCE 2867 AA; 331433 MW; 6E7D8CA71AFBFD3 CRC64;
Query Match 4.4%; Score 161; DB 1; Length 2867;
Best Local Similarity 18.4%; Pred. No. 3.6;
Matches 162; Conservative 147; Mismatches 302; Indels 268; Gaps 42;
Qy 9 LNESESSQGLGYFSDLNFPQVMTSTTGDLSIPSELENIPSENQYFQSAIWS-- 66
Db 927 LNELESG-----FKSLN-----GSASTTNKQENIRKNDITKSLNPAKN 966
Qy 67 -----GFTK-----VKSDFTYFATSDNHNVTMVDQF-----VINKASNSK 105
Db 967 SSSSKLSLENIIRKADLIKLDQHT--QEIEKHT--FIENEENGLSVIKKEKNRVE 1022
Qy 106 IRLEKGRLYQIK-----IOYQR-----ENPTEKGLDFKLYWTDG----- 139
Db 1023 SDMSSELIKQNTKYNALIEYNNKSDRPNFGDDETNLEELDDFKQCOQAQIKLTTN 1082
Qy 140 -----ONKKEVISSDN-----LQPELKQKSSNSRKRGRSTSAFTVDPDRD 179
Db 1083 YNVLDNGINVIIEKEHEKVIILSENHITKDKKINEKIQNVNSLNMETKGLGLKINED 1142
Qy 180 NDIPIPSL---EVEGY-----TVDKNKTFLSPWISNTHKKGLTKYKSSPEKWS 227
Db 1143 IKNSRDTTIKSKIQEFKKVQTTFGSDIVANKK-----IDAKKHEVDNKBDFKEXK 1196
Qy 228 TAS--DPYSDFEKVTGRI-----DKNVSPEARHPLVAAYPIVHVD----- 265
Db 1197 DTSPDEKKKGIKAYEKGNGNTLKEKMDDEKNIKEVEEBAQI--QYKRIEFDHVNLMND 1255
Qy 266 -----MENILISKNE--DQSTQNTDSEFTTISKNTSTSTRTSTNSNST--VAIDHS 313

Db 458 TNDISVEVEKEEEEEEEEN-----STFSKVKK-----ENVT---GEQEA 495
Qy 97 I--NKASNSKIRLEKGRLYQIKIQYQRENTEKGLDFKLYWTDSONKKEVISSDNILQLP 154
Db 456 VRNEVSTCEESTSKGE---EIMGDEKSEAGEKSSIIEISGSANSKIKONLVLE 551
Qy 155 ELKQKSSNSRKRGRSTSAFTVDPDRNDG-----IPDSLEV----- 189
Db 552 DEAEAPTQENKFTVEVGSIDIPAPRDDEIVEAVEKNIIPEDELEVAKEDEQEGVQLDE 611
Qy 190 -----EGVTVDKNTKTFSPWISNIHKKGL-----TKY 219
Db 612 PVKAMKDKIARGAESISEDKKQEGTAB--LSNEKAKKEVDSTABSAEGVEVSKT 670
Qy 220 KSSPE--KWTASDPYSFQKVTGRIDKNVSEARHPLVAAYPIVHVMENIILSKNEDQ 277
Db 671 PESPRVKRCTSGRP--EDLQ--INERDPEVLKEDVRVDPEDVKPIATTIEN---SEBEDP 725
Qy 278 STQNTDSTRITSKNTSRTSTNSNSTVAIDH-----SLSLAGER--TWAETMGLNT 330
Db 726 KSORQVIST---EQAEITQKMGVGGTTFKBEKPRFEITQEGDKITGKTQNTNHEG 781
Qy 331 ADTARLANIRYNTGTAPIYVNLFTSLVLGKQNTLATIKADENQLSQILAPNNYYPK 390
Db 782 EATEAASSENSKASDVGTAEKY--IEPSESV--KKDTEDEAVENSEKTEFI-----K 830
Qy 391 NLAPTA--LNAOKDASSTPTWYNGQLEL-----EKTKQLRLDITDQVYGNI--- 435
Db 831 VKAELENLAPKAEAVTAELNKENEDVEVDTEDEAVENSEKTEFIKVKAB--LGNLDAP 888
Qy 436 -----ATYFNGVRVVDVTSNWSVLPQIQETTARIIFNG-----K 472
Db 889 KEAVTAELNKENEDVEVAATSK--EDIETKCEPAETPIEDTCTCAEVSKDAEAVTKE 947
Qy 473 DLNLVERRTAA-----VNPSDPL----- 490
Db 948 DENMENSITAEALKQVGTGQEDIDNIDSEFORTVELPELEKQDKDKNGKEDKELEVEET 1007
Qy 491 --ETTKPDMTLKALKIAPGFNEPQNGNLQYQKGDITEFDNF-----FD 531
Db 1008 EKETSPLDLVVEE-----NITEKNEIKQEEBVEYSQLDNFNETESISKEAPNBNBNGFED 1061
Qy 532 QOTSQNIKNQLAB-----LNATNTYTVLDKTKLNAKKNILIRDKRHYDRNNIATVAGDE 585
Db 1062 QSTRENPKKASADDIPKOILDDETNEFLQKLVDDSELNALLQ-----SLDAK 1110
Qy 586 SVVKEAAREVINSSTGLLNTDKDIRKILSGYVIEIDTEGLKEVINDRYDMLNMISSLR 645
Db 1111 STQC-----TTEQSKKNDKP----- 1126
Qy 646 QDGKTFIDPKYNDKLPYI 665
Db 1127 QDVITTSITRKLNKEPVI 1146
RESULT 10
RBP2_PLAVB STANDARD; PRT; 2867 AA.
AC Q00789; Q9NZM3;
DT 01-APR-1993 (Rel. 25, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Reticulocyte binding protein 2 precursor (PVRBP-2).
GN RBP-2 OR RBP2.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A., AND REVISIONS TO 2438-2439.
RX MEDLINE=20299192; PubMed=10838229;
RA Galinski M.R., Xu M., Barnwell J.W.;
RT "Plasmodium vivax reticulocyte binding protein-2 (PVRBP-2) shares

Db 1256 EVEKSKIVMEKIELYKKEIDEIKQKTE---YKQDTSNFYVYEQVNSATQSKAKIEQF 1311
Qy 314 LSLAGERTWAETMGLTADATARNANIRVNTGTAPIYVNLPTTSLVGLKQKQTLATIKAD 373
Db 1312 INIA-----TTKGTSTSQ-----DINELESIKEEVHKH--LQLVQOE 1348
Qy 374 ENQL-----SQILAPNNYPSKRLAPIALNAQAQDASSTPTTMNNOFLLEKTKQLRLDTD 429
Db 1349 SNSMEEMRKQILSMKDLLLNNSETIA---KEISNN--TONALGFRENKTKLNK--TD 1400
Qy 430 QYVGNATYFNGRV---RVDTGNSWSEVLQIQ--ETTAIIIFNGKDLNLRERRIAAV 484
Db 1401 ELLQRYAAM--IBEAKAHKNIDIALEDQAIDTEVSKIEQINREIMNKD--EIKSYLSEI 1457
Qy 485 NPSDPLETTKPMWTAKKALIAF-----GFNPNGNLOVQGGKDIPEPDFN 529
Db 1458 KEYKDKCTTEISNKGKDKIFLEKFKPNESNSKNVKNINEINIRNSQYLKDIB-- 1515
Qy 530 FPOQTSQNTQKQALNATNIVTVLDKIKLNKAMILIRDRKFRHYDRNIAVGADESVMK 589
Db 1516 -DAEQKASTKVELFKHETHTISNIFKESI---LGVETKSQK-----KINKAEDIMK 1563
Qy 590 --EAHREVINSTEGLLNIDKIDIRKILSGYIVIEIDTEGLKEVIND----- 634
Db 1564 EIERHNSETQTVKGFQENLANK-----LNEPHNYDNADELNDKSTNAKVLIETN 1614
Qy 635 ----RYDMLNISLRQDG-KTFIDFKYNDKULPLYSN-----PNYKVNYYA 676
Db 1615 LESVKHNLSEITNIKQGGKIKYKARDIMQKIKATSENTAEKTBKVKDDQSNVYNLQ 1674
Qy 677 VTKENTIIIPSENG---DTSNIGIKILIFSKGVEIG 711
Db 1675 ITERNLIIVTEKNRLNGIDSTTINIEGALKESKGNVEIG 1713

RESULT 11

MDS3 YEAST
ID MDS3 YEAST STANDARD; PRT; 1487 AA.
AC P53094;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MDS3 protein (MCK1 dosage suppressor 3).
GN MDS3 OR YGI197W OR G1307.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_taxid=4932;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97197971; PubMed=9046087;
RA Coglievina M., Klima R., Bertani I., Delneri D., Zaccaria P.,
RA Bruschi C.V.,
RT Sequencing of a 40.5 kb fragment located on the left arm of
RT chromosome VII from Saccharomyces cerevisiae.;
RL Yeast 13:55-64(1997).
RN [2]
CHARACTERIZATION.
RP Li M.B., Neigeborn L.;
RA Unpublished observations (XX-1997).
RL
CC -!- FUNCTION: NOT KNOWN; NEGATIVE REGULATOR OF EARLY MEIOTIC GENE
CC EXPRESSION.
CC -!- SIMILARITY: CONTAINS 3 KELCH REPEATS.
CC -!- SIMILARITY: TO YEAST YER132C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; X91837; CAA62947.1; -;
DR EMBL; Z72719; CAA96909.1; -;
DR PIR; S62048; S62048.
DR Germline; 141245; -;
DR TRANSFAC; T03484; -;
DR SGD; S0003165; MDS3.
DR GO; GO:0005737; Cyttoplasm; IDA.
DR GO; GO:0030435; Psporation; IGI.
DR InterPro; IPR006652; Kelch_rep.
DR Pfam; PF01344; Kelch; 2.
KW Kelch repeat; Repeat; Meiosis.
FT REPEAT 171 226 KELCH 1.
FT REPEAT 234 287 KELCH 2.
FT REPEAT 371 419 KELCH 3.
SQ SEQUENCE 1487 AA; 167073 MW; 768AEPBAB796E447 CRC64;

Query Match 4.4%; Score 160.5; DB 1; Length 1487;
Best Local Similarity 19.3%; Pred. No. 1.6;
Matches 171; Conservative 137; Mismatches 271; Indels 305; Gaps 47;

Qy 2 VKQENLLNESBSSSQGLLYYFSDLNFOAPMVVTSS-----TTGDLISIPSSSELENIPS 55
Db 610 VREGEL-----SSGSLDNYF-EKNF--PIPARTSVSEAQNTQOVANADAKANTPS 660
Qy 56 ENQYFOSAIMSGFIKVKSGDEYTFATSDNHYTMV-----DQOEVIN--KASNSKI 106
Db 661 -----TSDEPSSSSSLLYSTPHYQRNDEBDDPVPSPKPVSKNSI 703
Qy 107 RLEKGLYQ--IKIQORENPTEKGLDLYWTD-----ONKKEVSSDNLQPELKQKSS 161
Db 704 -----YRPIKTESSTTSSNGMIFRVPFKEAAVTNTALELSLSQELSRRS 756
Qy 162 N-----SRKRTSAGPTVPRDNDGIP-----DSLEVEGYTVDVKNKT 201
Db 757 SLMSIPSGELLRSSISEAEHQRASHPLTSSPLFEDSGTFCGKQLQLOQHT--IQPHN 814
Qy 202 FLSP-----WISNIHEKGLTKYKSPKKNSTASPYDFEKVGTGIDKNVSP 249
Db 815 HLSFRFRSRSARSSISYVSSSSDRRG-----NSISSRST-SDSFGT-PVLGVNLVPLPP 867
Qy 250 EARHPLVAAYPIVHVDMENIILSKNDQ--STQNTDSETRITISKNTSRTHTNSN-- 304
Db 868 QREP-----NEPPPCPAMSGSNTR--RSNTLDYVHNSKASFP 906
Qy 305 -----SSTVAIDHSLAGERTWAETMGLTADATARNANIRVNTG--TAPIY 351
Db 907 SRRSSRHGRSSPTETENAFS--ATPRA-----SLDQMLGKSLKEGSTSQYTQPRM 957
Qy 352 NVLPTTSLVLGKQTLATIKADEN-----QLSQILAPNNY- 387
Db 958 NSFPKA-----NETIQTSSNNHRSQSVTSNTDPSLQSNFALEPLELTPRSLYM 1011
Qy 388 --PSKN-----LAPIALNAQAQDASSTPTTMNNOFLLE-----EKT 421
Db 1012 PWTSTVRAFAEPFYTGQVNSKWLAPVALDLVAKIYEIPLLYKLIILEVLYSLAKKE 1071
Qy 422 KQLRLDQVYVGNATYF-----ENGRVVDGTSNMSEVLQIQETARTIIFNG-KD 473
Db 1072 ESSLICTSLMETFRKTLNSYKGBDEKTYLTSDNTYQELLK--LKVSLENIDNGYVD 1129
Qy 474 LNIIVERRIAAVNPS-----DPLETTKPD----- 496
Db 1130 PDLRLKQSRQAQSSSTQESSGSANGEKATATGAGSLETSTNTVPTVAGGPRDSHNSVGSIG 1189
Qy 497 ---MTLKEALKIAFGNPNGL-----QYQOKDITEPDFDFDOOTSQ 537
Db 1190 PNMNMIQGRRTSGFS--PRVVKSSLSKEIDPKTFYEYEPKESKSDDDNDQDTNIG 1248
Qy 538 IKN-QLAELN-----ATNIYT-----VLDKIKLNKAMILIRDRKFRHYDRN 577
Db 1249 SFNLHLFDMYGISISSSTNSISSSSDLESKEQEQLDLETEREDSAEILDARF---RN 1305

```
QY 578 NIAVGADSVVKEAHREVINSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRYD 637
Db 1306 K-----EDDKYTKD-----ISNDKXNYLPHKXNKAKEG--KETRDVREBEEZFDGLG 1354
QY 638 MNISLSLQDGTFTDFEKKYNDKL-PLYISN--PNYKVNYYAVT 678
Db 1355 MSLNKKIKREAR-HVD--KVDDSDVDFPKSSAFQSPRAYGST 1395

RESULT 12
RBP1_PLAYE
ID RBP1_PLAYE STANDARD; PRT; 2869 AA.
AC Q00758;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reticulocyte binding protein 1 precursor.
GN RBP1.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites."
RL Cell 63:1213-1226(1992).
CC -!- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to
human reticulocyte cells.
CC -!- SUBUNIT: Homodimer (Potential).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
Db 2109 EMNTYNSIT-----QLEGIIVVSAGESKEDIKLEERSNEWRNISEKIS 2152
QY 295 TSRTHTSNSNSTVAIDHSLSLAGERTWATMGLMTADPARLNANIRYV-----NTGT 347
Db 2153 TIDSKVIEMNST---IDELYKL-GKNCQAHWISL-ISTANMKTSKKLIMINKENTSEK 2207
QY 348 APIYVNLPTSLVLGKNQTLATII---KADENQLSQILAPNNYPPSKNLA---PIALNAOK 401
Db 2208 CVDY-ISKDSSSTDGIVETLKGFGSKLTFSSASEIVQVADTY-SVNFAPKHESLNAR 2265
QY 402 DASSTPITMYNQFLEL-----EKTQKRLDLDQVYGNIAIYNFNGRVRVDT 449
Db 2266 DIKKELYLPHQNSDISIVEGGVQNMALYDLKNEEKREMDLYRNISETKLQMEHSTDV 2325
QY 450 GSNWSEVLPOIQETTARIIFNGKOLNLVRRRTAAVNPSPDLETTPKPDMTLKALKI---- 505
Db 2326 FKPMIELHKGMNET-----NNKSLLEKKEKKLSVN--DHMSMEABM-INKGLKYTPES 2376
QY 506 AFGFNEPNGNLQYQGDITFDFPNF--DQQTQNIKNQLAEL-NATNIYTVLIDKIKLNK 562
Db 2377 VQNNINIVSIEAEVKTLEIDRDYGDNYQIVVEHKKQFSILDRTN--ALMDDIEIFKK 2434
QY 563 MNILIRDKRFHY---DRNNIAGADESVVKEAHREVINST--EGLLI--DKDIRKI 614
Db 2435 EN-----NYNLMVNTTETIHRVNDYIEKTNKLVOAKTYEQILENIKQNDMLQNI 2486
QY 615 LSGYIVIEIDTEGLKE---VINDRYDM-----LNISSLRQDGKTFDFKKYNDKPL 663
Db 2487 FLKKVSIIEYFENVKVKESILNDLYQERLLAKIGELHDEIKENVTEILSSVEIQKMEM 2546
QY 664 YISN-----PNYKVNYYAVTKNTIINPS-----ENGDTSTNGIKKILIFSKKG 707
Db 2547 MSKNLLEKSKWMNY-TSYELEREANEINRDAKQIKDDDTILNSVLEAAI-QKRG 2600

RESULT 13
BXEN_CLOBU
ID BXEN_CLOBU STANDARD; PRT; 1162 AA.
AC Q06366;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Botulinum neurotoxin type E, nontoxic component.
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BL6340.
RX MEDLINE=93360835; PubMed=8355622;
RA Fujii N., Kimura K., Yokosawa N., Oguma K., Yashiki T.,
RA Takeshi K., Ohyama T., Isogai E., Isogai H.;
RT "Similarity in nucleotide sequence of the gene encoding nontoxic
component of botulinum toxin produced by toxigenic Clostridium
butyricum strain BL6340 and Clostridium botulinum type E strain
RT Mashike";
RL Microbiol. Immunol. 37:395-398(1993).
CC -!- FUNCTION: THE NONTOXIC COMPONENT IS NECESSARY TO MAINTAIN
TOXICITY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
DR InterPro; IPR000395; Peptidase_M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
KW Neurotoxin.
SQ SEQUENCE 1162 AA; 136829 MW; C8629BE932DA79E4 CRC64;

Query Match 4.4%; Score 159.5; DB 1; Length 1162;
Best Local Similarity 21.5%; Pred. No. 1.3;
Matches 173; Conservative 111; Mismatches 288; Indels 231; Gaps 43;

QY 2 VKQENRLNESSSOGLGYFSDLNFOAPWVTSITGDLSPSS-----ELENI 53
Db 372 INQNTVLMKSNFYGDGLG---NDNFVNTIIIPYLNYSINYSKLDNNVIEEIKI 428
QY 54 PSENO-----YFQSAIWSGFI---KYKSDXYTFATSAD--NHVTWVDDQEVINKASNSN 104
Db 429 PPIDEDDIYPRKNA--DTFIPVNIITAKAINTTTPVNVYLAQAMIDSDINLSSDFL 486
QY 105 KIRLEKGL-----YQIKIOYORENPEK-----GLDFPLYWTDQNKKEVIS- 147
Db 487 KVISKGLSVYFLANTMDYLFYKDYKPIDIDTKKYKWLKAIFRNSLDITQETISNQ 546
QY 148 -SDNQLPELQ-----KSSNRKRKSTSGA-----TVPDRDNDGIPDSLEVEG 191
Db 547 FGDTKIIPWIGRALNLTNTNSFVEEFKULGPIFLINKKENITIPKIDRIPSSM---- 602
QY 192 YIVDVKNKETFSPWISNIEHKG--LTK-YKSSPEKSTASDPYSDFEKVTGRIDKQVS 248
Db 603 LNFSPKD-----USEDNFYKRNFFYKIIYFNLDQWMT--QYYSQYFDLICKASKVL 656
QY 249 PEARHPLVAAPIVHVDMEIILSKNEDOSTQNTD--SETRTISKTSTSRHTSNSNST 307
Db 657 --AQEKLIK-----LIQRLYLMENSISNTLILNLTNTLRLDISNQSQ 703
QY 308 VAIDHLSLAGETWAEWTGLTADTARLANI--RYVNTGAPYVNVLPYTSVLVGLKQK 365
Db 704 IAINNDKPF-----FNNAMCVFENNIYKPF--SFMEQCINKNKST-----KEF 747
QY 366 TLATIKADENQLSQILAPNYPKSNLAPIALNAQKASSTPTVMYVNFQLELEKTKOLR 425
Db 748 ILKCTININETERKSHLMQNSP-----SNLQFD-----FLDIQNMKNL- 784
QY 426 LTDQVYGNIAVYFENGVRVDTSNNGEVLVPIQETARI--FNGKDLNLVERRIAA 483
Db 785 -----FNLYTELLIKEQTSPEYLSLYAFQEQDNNVIGDTSQKN-TLVE----- 826
QY 484 VNPSDELTETKPDMLKEALKIAGFNEBNGNLQYQKDIETFDNFQQTQSONIK--NQ 541
Db 827 -YPKD-----IGLVYGIN--NNAIHLTG-----ANQNIKFTND 856
QY 542 LAELNATN---IYTVLDKIKLAKNKNILIRK-----RFHYDRNRIAV-----GADES 586
Db 857 YFENGLTNFYSIYFMLRNLTQNTIKSKLIGSKEDNCGWEIYFENGLVFNIDSNGNEKN 916
QY 587 VV-----KEAREV--INSSTEGLLNID-----KQIRKILSGYIYEIED----- 624
Db 917 IYLSNISKSWHYIVISINRLKQQLIFIDNLTILVANEDIKELNTYSDIISLSDNNNV 976
QY 625 -TEGL-----REVINDRYDMLNLSISLRQDGKTFIDFKYNDKPLXYISNPNYKVV 674
Db 977 YIEGLSVLNTINSNLTDFSDLNNSVIRNFDEIL--QYNKTYELF-----NYVPE 1029
QY 675 YAVTK--ENTINPENGDTSN 695
Db 1030 IAINKIEQNNIYLSNNNSLN 1052

RESULT 14
ID SIR4 YEAST STANDARD; PRT; 1358 AA.
AC P11978;
DT 01-OCT-1989 (Rel. 12, Created)
```



```
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M14628; AAA33227.1; -.
DR FIR; A26655; A26655.
DR PDB; 1MMA; 03-DEC-97.
DR PDB; 1MMD; 17-AUG-96.
DR PDB; 1MMG; 03-DEC-97.
DR PDB; 1MMN; 03-DEC-97.
DR PDB; 1MND; 17-AUG-96.
DR PDB; 1MNE; 17-AUG-96.
DR PDB; 1VOM; 23-DEC-96.
DR PDB; 1LVK; 28-JAN-98.
DR PDB; 1DOX; 20-DEC-00.
DR PDB; 1DOY; 20-DEC-00.
DR PDB; 1DOZ; 20-DEC-00.
DR PDB; 1DIA; 20-DEC-00.
DR PDB; 1DIB; 20-DEC-00.
DR PDB; 1DIC; 20-DEC-00.
DR PDB; 1FNV; 20-DEC-00.
DR PDB; 1FWM; 20-DEC-00.
DR PDB; 1G8X; 17-JAN-01.
DR PDB; 1JWY; 07-NOV-01.
DR PDB; 1JX2; 07-NOV-01.
DR DictyBase; DDB0002015; mhca.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR008989; Myosin_S1_N.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00663; myosin head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PSS0096; IQ; 1.
DR Calmodulin-binding; ATP-binding; 3D-structure;
KW Myosin; Coiled coil; Actin-binding; Phosphorylation.
KW Calmodulin-binding; Methylation; MYOSIN HEAD-LIKE.
FT DOMAIN 1 761
FT DOMAIN 762 791
FT NP_BIND 817 2116
FT NP_BIND 179 186
FT DOMAIN 638 660
FT DOMAIN 738 752
FT MOD_RES 130 130
FT MOD_RES 1823 1823
FT MOD_RES 1833 1833
FT MOD_RES 2029 2029
FT TURN 3 5
FT TURN 7 8
FT TURN 10 15
FT TURN 25 28
FT STRAND 34 37
FT TURN 41 42
FT STRAND 48 55
FT STRAND 59 63
FT STRAND 69 73
FT STRAND 74 76
FT STRAND 78 79
FT TURN 83 85
FT TURN 86 87
FT STRAND 90 90
FT TURN 91 93
FT TURN 99 110
FT TURN 111 113
FT STRAND 116 119
FT TURN 120 121
FT STRAND 122 126
FT TURN 137 142
FT TURN 143 145
FT TURN 148 150
FT TURN 155 169
FT TURN 170 170
or send an email to license@sib-sib.ch).
-----
FT STRAND 173 179
FT TURN 181 182
FT HELIX 185 200
FT HELIX 210 226
FT STRAND 227 228
FT TURN 234 235
FT STRAND 236 237
FT STRAND 240 247
FT TURN 249 250
FT STRAND 253 261
FT HELIX 265 268
FT TURN 269 269
FT TURN 273 274
FT STRAND 278 278
FT HELIX 279 287
FT HELIX 290 296
FT TURN 297 297
FT TURN 301 303
FT TURN 305 307
FT TURN 316 317
FT TURN 320 334
FT TURN 335 335
FT HELIX 338 355
FT TURN 356 356
FT STRAND 360 360
FT STRAND 368 368
FT HELIX 373 382
Query Match 4.4%; Score 159.5; DB 1; Length 2116;
Best local similarity 19.3%; Pred. No. 2.8;
Matches 131; Conservative 116; Mismatches 232; Indels 201; Gaps 33;
86 HVTWVDDQEVINKASNSKIRLEKGRVQIKIQYORENPTEKGLDFKLYWTDSONKKE- 144
1294 HVNEQLEEEKKQKESNEKKVLDLEK-EVSELKDQIEEFVASKAV-----TEAKNKES 1346
145 -----VSSDNQLQPELKQKSSNSRKRSTSGPTVPDRNDGIPDSLEVEGYT 193
1347 ELDEIKHQYADVSSRDKSVQELKTLOAKNEELRNTA-----BEAQGLDRAERS--- 1396
194 VDVNKRKTF-LSPWISNIHEKGLTKYKSPKSWSTADSPYSFEKVTGRID--KNVSPE 250
1397 ---KKKAEFDLEEAIVKNEEE---TAKVKAELAMKKAEE--IDYRTKSELDDAKNVSE 1448
251 ARHPLVAAPYIVHDVMDENIILSKNEPQS-----TQNTDSETRTISKNTSTSRTHSNS 305
1449 Q-----YVQIKRL---NEELSELRSVLEBADERCNSAIKAKKTAESALESKD 1493
306 STVAIDHSLSLAGERTWAETWGLNTADTARL-----NANIRVYNTGTAPIYVLPITSL 359
1494 EIDAANNAKAKAEK--SKLEVRVAELLESLEDKSGTVNVEFIRKDAEIDDLRLDL- 1550
360 VLGNQTLATIKADENQLSQILAPNNYPSKNLAPIALNAQKSDASSTPITMNYNQFLELE 419
1551 ----RETESRIKSDK-----KNTRK-----QFADLE 1574
420 KTKQLRLDTQVYGNATYFNGRVVDTGSKWSEVLPOIQ-ETTARI-IFNGK---DL 474
1575 -----AKVEEAQREVVTID-----RLKKLESIDILSTQDITETSRIKESKKLEQ 1624
475 NLVERRIAANVPSDFLETTKPKDMLKEALKIAFGFNEPNGNLQVQGDITFDFNPDQ- 533
1625 TLAERRAEAGSSKAAD-----EIRKQVW-----QEVDLRAQLDSER 1663
534 -----TSQNTKQKQALNAELNATNYVLDKIKLNAMNILLRDKRPHYDRN---NIAVGADE 585
1664 AALNASEKKIKSLVAE-----VDEVKEQLEDEILAKDKLVKAKRALEVEEVRDQ 1714
586 SVVKEAHRVINSSTEGLLNIDKIRKILSGYIETE-----DTEGLKEV 631
1715 LEEBEDSRSELEDSKRLLTTEVE--DIRK---KYDAEVEQNTKLDEAKKLTDDVDTLKKQ 1770
632 INDYDMLNISS-----LRQDKTFI-----DFKKYNDKLPYISNPYK 671
QY
```


Db 1771 LEDEKKKLNESERAKRLESENEDEFLAKLDAEVKNRSRAEKORKKYEKDL-----KOTKYK 1826
QY 672 VNVYAVTKENTIIINPSENGD 691
Db 1827 LNDERATKTQTEIGAKLED 1846

Search completed: May 3, 2004, 19:37:07
Job time : 10.9844 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 33.4955 Seconds
(without alignments)
6697.424 Million cell updates/sec

Title: US-09-848-909A-16
Perfect score: 3648
Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKKILFSKKGYEIG 711

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rviris.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3535	96.9	741	2 Q937W2	Q937W2 bacillus an
2	3535	96.9	743	2 Q937W3	Q937W3 bacillus an
3	859.5	23.6	876	2 Q32739	Q32739 clostridium
4	857.5	23.5	876	2 Q9KH41	Q9KH41 clostridium
5	849.5	23.3	879	2 Q06498	Q06498 clostridium
6	849	23.3	875	2 Q45221	Q45221 clostridium
7	792	21.7	721	2 Q86171	Q86171 clostridium
8	705	19.3	775	2 Q844J8	Q844J8 bacillus th
9	296.5	8.1	225	2 Q8XYK2	Q8XYK2 bacillus th
10	218.5	6.0	1881	16 Q8RGK2	Q8RGK2 fusobacteri
11	205.5	5.6	675	16 Q897H6	Q897H6 clostridium
12	203.5	5.6	1301	5 Q8SKS5	Q8SKS5 plasmodium
13	200	5.5	4688	16 Q9PQ08	Q9PQ08 ureaplasma
14	199.5	5.5	2849	5 Q8HY44	Q8HY44 plasmodium
15	199.5	5.5	3468	5 Q81104	Q81104 plasmodium
16	199	5.5	1404	2 Q8RJN9	Q8RJN9 mycoplasma

17	198.5	5.4	1301	5 Q8IHQ2	Q8IHQ2 plasmodium
18	197	5.4	1302	2 Q49547	Q49547 mycoplasma
19	196.5	5.4	1904	5 Q8IKS2	Q8IKS2 plasmodium
20	196	5.4	3248	5 Q8ICP9	Q8ICP9 plasmodium
21	194	5.3	1859	5 Q8IC27	Q8IC27 plasmodium
22	194	5.3	2792	5 Q814R2	Q814R2 plasmodium
23	192.5	5.3	2529	16 Q25579	Q25579 helicobacte
24	192.5	5.3	2771	5 Q26216	Q26216 plasmodium
25	192.5	5.3	5767	5 Q81525	Q81525 plasmodium
26	192	5.3	1833	5 Q81A25	Q81A25 plasmodium
27	192	5.3	2940	5 Q81HP9	Q81HP9 plasmodium
28	192	5.3	9904	16 Q8NWQ6	Q8NWQ6 staphylococ
29	190.5	5.2	1263	5 Q81K49	Q81K49 plasmodium
30	190.5	5.2	1722	5 Q81118	Q81118 plasmodium
31	190	5.2	752	16 Q98PL9	Q98PL9 mycoplasma
32	189.5	5.2	2577	5 Q81531	Q81531 plasmodium
33	188.5	5.2	4095	2 Q7WZ13	Q7WZ13 haemophilus
34	188	5.2	1425	5 Q81K08	Q81K08 plasmodium
35	187	5.1	2269	5 Q26223	Q26223 plasmodium
36	187	5.1	6713	16 Q99U54	Q99U54 staphylococ
37	187	5.1	6713	16 Q931R6	Q931R6 staphylococ
38	186.5	5.1	990	5 Q815C6	Q815C6 plasmodium
39	186	5.1	2747	5 Q9BJX9	Q9BJX9 plasmodium
40	185.5	5.1	2616	5 Q81IG1	Q81IG1 plasmodium
41	185	5.1	1936	5 Q815A6	Q815A6 plasmodium
42	184.5	5.1	2578	5 Q81JP9	Q81JP9 plasmodium
43	184	5.0	1151	5 Q81LU0	Q81LU0 plasmodium
44	184	5.0	2072	5 Q81I32	Q81I32 plasmodium
45	184	5.0	2752	5 Q9BJY0	Q9BJY0 plasmodium

ALIGNMENTS

RESULT 1

Q937W2 ID Q937W2 PRELIMINARY; PRT; 741 AA.
AC Q937W2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Rag protein (Fragment).
GN PAG.
OS Bacillus anthracis.
OG Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ferrara;
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
RA Fasanella A., Francia M., Ciuchini F.,
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus anthracis vaccine strain 'Carbosap'.
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ413937; CAC93935.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxins.
DR Pfam; PF03495; B:binary tox8; I.
DR PRINTS; PR01391; BINARYTOXINE.
KW Plasmid.
FT NON_TER 1
FT NON_TER 741
SQ SEQUENCE 741 AA; 83153 MW; C7F95820B73065C0 CRC64;

Query Match 96.9%; Score 3535; DB 2; Length 741;
Best Local Similarity 96.0%; Pred. No. 2.8e-160;
Matches 693; Conservative 1; Mismatches 4; Indels 24; Gaps 1;

Cy 1 EVKQENLLNESSSQGLLYFFDLNFPAPMVVTSSTGTGLSPSSLENIPISENQYF 60

Db 20 EVKQENRLNSESSESSQGLLYYFSDLNFCAPMVVTSSTGDLSPSSLENIIPSENQYF 79
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 80 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 139
QY 121 QRENTEKGLDFKLYWTDSONKEVSSDNQLPELKOKSSNSRKRSTAGTPVDRDN 180
Db 140 QRENTEKGLDFKLYWTDSONKEVSSDNQLPELKOKSSNSRKRSTAGTPVDRDN 199
QY 181 DGIPLSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240
Db 200 DGIPLSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 259
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
Db 260 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 319
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 336
Db 320 SEVHGNAEYHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 379
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 396
Db 380 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 439
QY 397 LNAQKDSASPTITMNYNQFLEKTKQLRLDQVYGNIAATYVFNENGRVVDTSNWSSEV 456
Db 440 LNAQDFFSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVFNENGRVVDTSNWSSEV 499
QY 457 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 516
Db 500 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 559
QY 517 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 576
Db 560 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 619
QY 577 NNIAGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVEIDTEGLKEVINDRY 636
Db 620 NNIAGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVEIDTEGLKEVINDRY 679
QY 637 DMLNSSLRQDGKTFIDFKYNDKPLFLYISNPNYKNVYAVTKENTIINPSENGDTSTNG 696
Db 680 DMLNSSLRQDGKTFIDFKYNDKPLFLYISNPNYKNVYAVTKENTIINPSENGDTSTNG 739
QY 697 IK 698
Db 740 IK 741

RESULT 2

Q937W3 ID Q937W3 PRELIMINARY; PRT; 743 AA.
AC Q937W3; 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pag protein (Fragment).
GN PAG
OS Bacillus anthracis.
OG Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Carbosap.
RA Adone R.; Pasquali P., La Rosa G.; Marianelli C.; Muscillo M.,
RA Pasanella A.; Francia M.; Ciuchini F.;
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus anthracis vaccine strain 'Carbosap.'";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ413936; CAC93934.1; -.
DR GO: 0005576; C:extracellular; IEA.
DR GO: 0046821; C:extrachromosomal DNA; IEA.
DR GO: 0015070; F:toxin activity; IEA.
DR GO: 0009405; P:pathogenesis; IEA.
DR InterPro: IPR003896; Anthrax toxinB.
DR Pfam: PF03495; Binary toxB; I.
DR PRINTS: PR01391; BINARYTOXINB.
KW Plasmid.
KW NON_TER. 1 743
FT NON_TER 743 743
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5D85A6 CRC64;

Query Match. 96.9%; Score 3535; DB 2; Length 743;

Best Local Similarity 96.0%; Pred. No. 2.8e-160;

Matches 693; Conservative 1; Mismatches 4; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFCAPMVVTSSTGDLSPSSLENIIPSENQYF 60
Db 22 EVKQENRLNSESSESSQGLLYYFSDLNFCAPMVVTSSTGDLSPSSLENIIPSENQYF 81
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 82 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 141
QY 121 QRENTEKGLDFKLYWTDSONKEVSSDNQLPELKOKSSNSRKRSTAGTPVDRDN 180
Db 142 QRENTEKGLDFKLYWTDSONKEVSSDNQLPELKOKSSNSRKRSTAGTPVDRDN 201
QY 181 DGIPLSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240
Db 202 DGIPLSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 261
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
Db 262 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 321
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 336
Db 322 SEVHGNAEYHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 391
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 396
Db 382 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 441
QY 397 LNAQKDSASPTITMNYNQFLEKTKQLRLDQVYGNIAATYVFNENGRVVDTSNWSSEV 456
Db 442 LNAQDFFSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVFNENGRVVDTSNWSSEV 501
QY 457 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 516
Db 502 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 561
QY 517 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 576
Db 562 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 621
QY 577 NNIAGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVEIDTEGLKEVINDRY 636
Db 622 NNIAGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVEIDTEGLKEVINDRY 681
QY 637 DMLNSSLRQDGKTFIDFKYNDKPLFLYISNPNYKNVYAVTKENTIINPSENGDTSTNG 696
Db 682 DMLNSSLRQDGKTFIDFKYNDKPLFLYISNPNYKNVYAVTKENTIINPSENGDTSTNG 741
QY 697 IK 698
Db 742 IK 743

RESULT 3

O32739

ID O32739 PRELIMINARY; PRT; 876 AA.

AC O32739;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE ADP-ribosyltransferase.
GN CDTB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
RN (1) _____
RP SEQUENCE FROM N.A.
RC STRAIN=CD196;
RX MEDLINE=97230316; PubMed=9119480;
RA Perille S., Gilbert M., Bourlioux P., Corthier G., Popoff M.R.;
RT "Production of a complete binary toxin (actin-specific ADP-
ribosyltransferase) by Clostridium difficile CD196.";
RL Infect. Immun. 65:1402-1407(1997).
DR EMBL; L76081; AAB67305.1; -;
DR HSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxB; I.
DR PRINTS; PR01391; BINARYTOXINB.
KW Transferase.
SQ SEQUENCE 876 AA; 98797 MW; 255062D45C2B3B CRC64;
Query Match 23.6%; Score 859.5; DB 2; Length 876;
Best Local Similarity 32.1%; Pred. No. 5.7e-33;
Matches 251; Conservative 131; Mismatches 277; Indels 123; Gaps 31;
QY 4 QENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSELEN-IPSENQYFQS 62
DB 40 KKKEIVNEILPNGLMGYFSDHEFKDLKMAPIKDGNLFEEKKVDKLDKDKSDVKS 99
QY 63 AIWSGFIVKKSDEYTFATSDNHNVTMVDQEVINKASNKIRLEKGRLYQIKIQYOR 122
DB 100 IHWGRIIPSKDGEVTLSTDRDD-VLMQVNTESI---SNTLVNMMKKGKVKYRIQLD 155
QY 123 EN--PTEKGLDKPLVWTSQNKKEVISSDNLQLPKOKSSNKKRSTSGTPV 176
DB 156 KNLGSDNLSPNLW-ELDGMKKIIPENLFRDY-----SNIEKD-----DPFIPNNF 205
QY 177 -----DRDNDGIPDSLEVEGYTVDVKNKTFSPWLSNIHEKKGLTKYKS 222
DB 206 FDPKMSDWEDEDLTDNDNIPDSYERNGYTI---KDLIAVKWEDSPAE-QGYKYYVN 260
QY 223 PEKWTASDPYDFEKVTGRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQST--- 279
DB 261 YLESNTAGDPYTDYEKASGSPDKAITEARDPLVAAPYIVGVGMEKLIISTNEHASTDQ 320
QY 280 -----QNTDSETRTISKN-----TSTSRHTSNSSSTVAIDHSLSLAGERTWA 323
DB 321 KTVSRATNSKTESAGVNVGVQNGFTANVTNYSHTDNTAVQDS---NGE-SWN 376
QY 324 ETWGLNTADTARLANIRVYGTAPIYNNVLTTSVLGKQNTLATIKADENQSLAP 383
DB 377 TGLSINKGESAVINAVRYNGTAPMYKVPTTNVL-DGTLSTIKAEQNGNLS 435
QY 384 NNYPSKNLAPALNAQKDASTPTMYNGVLEKTKYQLRDLTDQYVGNIAFYENG 443
DB 436 GDTYPPKGLSPALNTMDQFSRLPIYNDQKLKLDAGKQIKLETQVSGNFGTKN-SSG 494
QY 444 RVRVDTGNWSVLPQIQTETARIIFNGKLNLEVERIAAANPSPDLETTKPDMLKEAL 503
DB 495 QI-VTEGNSWSDYISQDISASIIIDTEN-ESVRRVTKNQLQDPEDKT-PELTIGEAI 551
QY 504 KTAFGNPNGLQYQKGDITE--PDFNPDQOTSQNTKQIAELNATNIYTVLDKIKLNA 561
DB 552 EKAFGATKDGLLYFNDFIPDESCVELIFDNTANKIKDLSKLTSLDKKIYVNV----KLER 607

QY 562 KXNILIRDKRPHY---DRNNAVGADSVVKEAHREVINSSTEGL-----LINIDK 609
DB 608 GWNILIKTYTFNPDYNNYP--STWSNVNTNQDGLQGSANKLNGETKIKIPMSLKP 665
QY 610 DIRKILSGY-----IVEI-----EDTEGL-----KEVINDRYDMLNIS-SL 644
DB 666 YKRYVFSVSGDPLTNSIIVIRKAKBEKTDYLVPEQGYTKFSYFETTEKDOSSNIEITL 725
QY 645 RQDGKTFIDFKYND--KLPLVISPNYKV-----NVYAVTKENTIINPSENGDTSTN 695
DB 726 IGSSTYLDNLSTELNSTFELDEPEVKIPTDQEIIMDAHKIYFADLNFPS-TGNTYIN 784
QY 696 GI 697
DB 785 GM 786
RESULT 4
Q9KH41 PRELIMINARY; PRT; 876 AA.
ID Q9KH41
AC Q9KH41;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE CdtB.
GN CDTB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
RN (1) _____
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 20309;
RA Chang S.Y., Song K.P.;
RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain
RT CCUG 20309.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP271719; AAF81761.1; -;
DR HSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxB; I.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;
Query Match 23.5%; Score 857.5; DB 2; Length 876;
Best Local Similarity 32.0%; Pred. No. 7.1e-33;
Matches 250; Conservative 132; Mismatches 277; Indels 123; Gaps 31;
QY 4 QENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSELEN-IPSENQYFQS 62
DB 40 KKKEIVNEILPNGLMGYFSDHEFKDLKMAPIKDGNLFEEKKVDKLDKDKSDVKS 99
QY 63 AIWSGFIVKKSDEYTFATSDNHNVTMVDQEVINKASNKIRLEKGRLYQIKIQYOR 122
DB 100 IHWGRIIPSKDGEVTLSTDRDD-VLMQVNTESI---SNTLVNMMKKGKVKYRIQLD 155
QY 123 EN--PTEKGLDKPLVWTSQNKKEVISSDNLQLPKOKSSNKKRSTSGTPV--- 176
DB 156 KNLGSDNLSPNLW-ELDGMKKIIPENLFRDY-----SNIEKD-----DPFIPNNF 205
QY 177 -----DRDNDGIPDSLEVEGYTVDVKNKTFSPWLSNIHEKKGLTKYKS 222
DB 206 FDPKMSDWEDEDLTDNDNIPDSYERNGYTI---KDLIAVKWEDSPAE-QGYKYYVN 260
QY 223 PEKWTASDPYDFEKVTGRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQST--- 279
DB 261 YLESNTAGDPYTDYEKASGSPDKAITEARDPLVAAPYIVGVGMEKLIISTNEHASTDQ 320
QY 280 -----QNTDSETRTISKN-----TSTSRHTSNSSSTVAIDHSLSLAGERTWA 323

```

Db 321 KTVSRATNTSKTESNTAGVSVVGVQNGFTANVTNNYSHHTDNSTAVQDS---NGB-SWN 376
QY 324 ETMGSLNTADTARLNANIRVYNTGTAPIYVNLPTTSVLGKQNTLATIKADENQLSQILAP 383
Db 377 TGLSINKGESAYINANVRVYNTGTAPYKVTPTTNLVL-DGDTLSTIKAQENQ-CNNLSP 435
QY 384 NNYPSKNAIAPALNAQAQDASSTPITMNYNQPLEKTKQLRLDQVYGNATYVNFENG 443
Db 436 GDTYFKGLSPALNTMOPFSRLPIYNDQKLUDAGKQIKLETQVSGNFGTKN-SSG 494
QY 444 RVRVDTGSNWSVLPOIQTETARIIFNGKDLNLVERRIAAVNPSPDPLETTKPDMLKEAL 503
Db 495 QI-VTEGNSWSYIQQIDSISASIILDTEN-ESYERRVATKNLQDPEDKT-PELTIGEBAI 551
QY 504 KIATFNEBENGLOVQGRDITE--PDFNFDQOTSQNIKNQLAELNATNIYVLDKILNA 561
Db 552 EXAFATKQGLYFNPIIDPDESCVELIFDDNTANKIKDSUKTUSDKKIYVNV---KLER 607
QY 562 KNNILIRKRFHY--DRNNIAGVAGESVVKAEHREVINSSTEGT-----LLNIDK 609
Db 608 GNNILIKTPTFTNFDYNNYP--STWSNVNTTNKDGQSGANKLNGETKIKIPMSCLKP 665
QY 610 DIRKILSGY-----IVZL-----EDTEGL-----KEVINDRYDMLNIS-SL 644
Db 666 YKRYVFGSVKDPILSNSIIVIKAKEKRTDYLVEPQGYTKFSYEFETEKDSSNIEITL 725
QY 645 RQDGKTFIDFKYND--KLPLVISPNTVK-----VNVYAVTKENTIIINSENGDTSN 695
Db 726 IGSCTIYLDNLSTELNSPFEILDPEVKIPPTDQELIDAHKIYFADLNFPNS-IGNTVIN 784.
QY 696 GI 697
Db 785 GM 786

RESULT 5
O06498 PRELIMINARY; PRT; 879 AA.
AC O06498;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE sb component.
GN SBS.
OS Clostridium spiroforme.
OC Bacteria; Firmicutes; Mollicutes.
OX NCBI_TaxID=29348;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS246;
RA Gilbert M., Perelle S., Daube G., Popoff M.R.;
RT "Clostridium spiroforme toxin genes are related to C. perfringens iota
toxin genes but have a different genomic localization.";
RL Syst. Appl. Microbiol. 20:337-347(1997).
DR EMBL; X57969; CAA66612.1; -.
DR HSSP; P13423; 1ACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:apathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB; 1.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 879 AA; 98739 MW; 40685ACB8E05BA01 CRC64;

Query Match 23.3%; Score 849.5; DB 2; Length 879;
Best Local Similarity 28.1%; Pred. No. 1.7e-32;
Matches 252; Conservative 143; Mismatches 250; Indels 253; Gaps 32;

QY 4 QENRLNSESQGLGYFSDLNFAQPMVYTSSTTGDLSIPSSSELENIPSENQ-YFQS 62
Db 36 QEVETTNETVSSNGLMGVYFADHPKDLLEMAPKVGKELKEKVKELLTEENTNIKS 95

```

```

QY 63 AIMSGEIKYKSDXYTFATSDADNHVTWYDDQGEVINKASNSKIRLEKGLYQIKIYOR 122
Db 96 IRTWGRILIPSKDGEYTLSTDXON-VLMQINAB--GEIANTLKVNMKQGEYSIRIEIO- 150
QY 123 ENPTERGLDF-----KLYWTDSONKKEVISDNLQLPKQKSSN----- 162
Db 151 ----DKDIGVDDLSLSPKLYW-ELNGDKTLIPKQLFLRDYSKIDENDPFIKDNFFDLK 205
QY 163 --SRKKSRTSA-GPTVPDRDNDGIDPSLVEGEYTVVKNKRTFLSPWISNIEHKGLTKY 219
Db 206 LKRSARLASGDCDELDTDNDNIPDAYEKNGYTI-----KOSIAVKW-EDSFAQGYKY 260
QY 220 KSPFKWSTASDPYDFEKVTGRIDKNVSPFEARHPLVAAYPIVHVDMENIILSKHEDQST 279
Db 261 LSSYLESNTAGPYTDYQKASGDFKAIKAEARDPLVAAYPVVGVGMEKLIISTNEHAST 320
QY 280 QNTDSTRISKNTSTSRTHSTNSNSTVAID-----HSLSLA 317
Db 321 ----DQGTQVSRNTTNSKT---DANTAGVAINAYONGFTGSIITNYSHSTENSTAVQNS 373
QY 318 GERTWAETVGLNTADTARLNANIRVYNTGTAPIYVNLPTTSVLGKQNTLATIKADENQL 377
Db 374 NGESWNTLSINKGESAYINANVRVYNTGTAPYKVTPTTNLVL-DGDTLTIKQADNOI 432
QY 378 SQILAPNNYPSKNLAPIALNAQKODASSTPITMNYNQPLEKTKQLRLDQVYGNAT 437
Db 433 GNNLSNETYPKGLSPLALNTMDQFSSRLIPIYNDQKLKLAGKQIKLETQVSGN--- 489
QY 438 YNFENGKRVVDV--GSNWSVLPQIQTETARIIFN-GKDLNLVERRIAAVNPSPDLETTPK 495
Db 490 YGINKSQGQIITEGNSWSYISQIDSLSASIILDTGSD--VFERRVATKDSNPEDKT-P 546
QY 496 DMTLKEALKIAGFNENFNGNLQYQGRDITE--PDFNFDQOTSQNIKNQLAELNATNIYTV 553
Db 547 VLTIGEAIEKAFGATKNGEILYFNGMPIDESCVELIFDGTANTANLKERLNALNDKKIYV 606
QY 554 LDKIKLNAXNNILIRD-----IAVGADESUVKAEHREV-----INSSST 600
Db 723 NIEITTSQGTIFLDNLSTELNSTPEILKEPDKVPDQEI--DAHKIYADLSFNQST 781
QY 601 EGLLN-----IDKIRKILSGYIVIEI-DTEGLKEVINDRYDMLNMISSLRQDKGT- 650
Db 782 ANVYLDGLYFEPQTQNTKNEVLVDYIKYKVEATLEYSFGDKIGTKDKELRNYTGDSNQPKTN 841
QY 651 FIDFKYNDKPLIYISNPNY----KVVYAVTKENTIIINPSENGDTSNGIKKILIFS 704
Db 842 YVNFERSY-----FTSGENVMYPYKLRUYAITPEN-----KELLVLS 877

RESULT 6
Q46221 PRELIMINARY; PRT; 875 AA.
AC Q46221;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Iota toxin component Ib precursor.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 10748;
RC MEDLINE=94041637; PubMed=82255592;
RX

```


286 TRTSKNTSTSRHTSNTSSTVAIDHSLSLAG-----BRTWAET 325
287 TKSMKSTSHSST--NINTVGAEVSGSLQAGGIPVPSMSASANYSHTWNTSTVDVT 343
326 -----NGLTADARLANANIRYNTGTAPIYVNLPTSLVLGNQTLATIKADENQL 377
344 TBSFQGSINIGESAYINPIRYNTGTAPIYVNPITTTIVDK-OSVATIKQESLI 402
378 SOILAPNNYPSKNLAPIALNAQKASSTPTIMYNNQFLELEKTKQLRLDTPQVYGNAT 437
403 GOVNLPGGYPIIGEPMALNTMDQSSSLIPINYNQLKSIDNGTVMNSTSQFTGNFAK 462
438 YNFENGVRVDTGNSWSEVLPOIETTARII--FNGKDLNLVERIAAVNSDPLETTKP 495
463 YN-SGNLVT-D-GNNWGPYLGITKSTASLTLSFGQTTQVA--VWAPNFSDPEDKT-P 516
496 DMTLKEALKIARFENPNGLNLOVQGDIT--EFDFNFDQQTSONIKQLAELNATNTYT 552
517 KLTLEQALVKAFLEKKNKGFYHGLEISKNEKIQVFLDSNTNNFENQNLKTADKDIWH 576
553 VLDIKILNAKMNILIRDKRPHYDRNNAIV 581
577 CI--IKRN--MNILVKVITKENISSINI 601

RESULT 8
Q844J8
ID Q844J8 PRELIMINARY; PRT; 775 AA.
AC Q844J8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ViPlac.
GN ViPlac.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y., Chen J., Pang Y.;
RT "Cloning of vip1A(C) and vip2A(C) from Bacillus thuringiensis HD201.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY245547; A086514.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03435; B:toxinB; I.
DR PRINTS; PR0139; BINARYTOXINB.
DR SQU SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;

Query Match 19.3%; Score 705; DB 2; Length 775;
Best Local Similarity 28.9%; Pred. No. 1.1e-25;
Matches 220; Conservative 129; Mismatches 285; Indels 128; Gaps 29;

QY 4 QENRLNESSSQGLGYFSLNFCAPMVVTSSTGSLSPSSELEN--IPSENQYFQ 61
DB 42 QKNQ---QKMDRKGGLGYFKGDF-SNLTMAPTRDNTLYDQQTANKLDRKQOQY 97
QY 62 SAWSGFVKVKSDEVTFTASDNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIYO 121
DB 98 SIRWIGLQSKGKDEFTFNLSDEQALIIDGKLIISNKGKQVHLEKLVPIKIEVQ 157
QY 122 RENPTKGLD-----FKLYWTSQNKKEVIVSDNQLPBLKQKSN----- 162
DB 158 SD--TKFNIDSKTFKFKLFDKIDSONQSOQVKRDELNPFENFKGSEREFLAKASKTNPFM 215
QY 163 SRKRSSTAGTVPDRNDGIPDSLEVEGYVTVVKNKRTFLSPWISNIHEKGLTKYKSS 222
DB 216 QKMKRDIDED--TDTGDSIPDLWEENGYT--IQNK--VAVK--DDKFAQGYQYKYLSS 267
QY 223 PEKWTASDYSDFEKVGTGRIDKNVSPEARHPVAAAYPIVHVDMMENILSKNEDOSTQNT 282

268 PYQARTVGDPTDWEKAAGDIPKSNAAATRNPLVAAFPINVDMRKWILSKDNLNLS---- 323
283 DSETRITKNTSTSRHTSNTS-----NSSTVAID-----HSLSLAGERTWAETMG--- 327
324 -----NSAEHSSNYSYANSEGASIEAGFGPKGFGVGSANTYQH:ETVGSWDG 372
328 -----LNTADARLANANIRYNTGTAPIYVNLPTSLVLGNQTLATIKADENQLSQ 379
373 NSKSNTQEFNSASAGYILNANVHYNNVGTGGIDYDAQPTTSPIL-QDSTIATITAKSNATAL 431
380 ILAPNNYPSKNLAPIALNAQKASSTPTIMYNNQFLELEKTKQLRLDTPQVYGNATYN 439
432 SIFSGRYPASK-EGSLKTMDDFNHPTLNKPOLDAVLNNEVIKINTDQDGRYGIIG 490
440 FENGVRVDTGNSWSEVLPOIETTARIIFNGKDLNLVERIAAVNSDPLETTKPDML 499
491 VDG---KABICDRWSPTIIDEIKGRTASIIIDPADGKALETIAAKDYKNPEDKT-PSLTI 546
500 KEALKIARF--FNEPGLNLOVQGD---ITEFDEN-----PQQTSONIKQLAEL 545
547 KEGKIATPESISEDKGILFYEKNDGKVTYKQJSEENIMPYLDEDTSKFERQJSDG 606
546 NATNIVTVLTKILNAKMNILIRDKRPHYDRNNAIVGADSVVKEAHREVINSSTGLL- 604
607 SAKGLY---DIKLTQKMNITIR-----LATVTLGDDQFSAYPEWENATWSDKFGNLR 655
605 -----LNIDKDIR-----ILSGVIVE--IEDTEGLKEVI--NDRYDMLNLS 643
656 LGSLATPQESKYTIPKDKVKNYDITGYIKHDFTTDNESIGIVAFTKKDNFEMNMGT 715
644 LRQDGKTFIDFKKYNDKPLPIYISNPNYKVNVAVTKEINTIIN 685
716 SIFSQNSGGFEKFTIKTQ-NISS-DYILDSIQLMKRNDVN 755

RESULT 9
Q8KYK2
ID Q8KYK2 PRELIMINARY; PRT; 225 AA.
AC Q8KYK2
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protective antigen-related protein, (px01-111).
GN BXA0163.
OS Bacillus anthracis.
OX Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2012;
RX MEDLINE=22061436; PubMed=12004073;
RA Read T.D., Salzberg S.L., Pop M., Shumway M., Umayam L., Jiang L.,
RA Holtzapfel E., Busch J.D., Smith K.L., Schupp J.M., Solomon D.,
RA Keim P., Fraser C.M.;
RT "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in
RT Bacillus anthracis.";
RL Science 296:2028-2033(2002).
DR EMBL; AE011190; AAM26108.1; -
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 225 AA; 25402 MW; 2E121BB54295F9C8 CRC64;

Query Match 8.1%; Score 296.5; DB 2; Length 225;
Best Local Similarity 37.2%; Pred. No. 5.9e-07;
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;

QY 542 LAELNANNTVLDKIKLNKMNILIRDKRPHYDRNNAIVGADSVVKEAHREVINSSTE 601
DB 1 MESLGININYNALDRILKLNKMNILVRDP-YHYDNGNIVGVDDSYLNKAYKILNWSDD 59
QY 602 GLLINIDKDIRKILSGYVIEPTE-----GLKEVINDRYDMLNLSLRQDG 648

Db 60 GVSINLDEVDNQALSGYMLQIKKPSNHLTNSPVITTLAKDQSGVGLYRVL-----DG 113

Qy 649 KTFIDFKYNDKPLLYISNPNKYVAVYKFN-TIINPSENGDTSTNGIKKILISKKG 707

Db 114 TGFLDFNKFDENWRLV-DPGDDVYVAVTKEDFNVDRENGNIA-NKDKNTLVLSGKI 171

Qy 708 YEI 710

Db 172 KEI 174

RESULT 10

Q8RGK2 PRELIMINARY; PRT; 1881 AA.

ID Q8RGK2

AC Q8RGK2

DT 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Hemolysin.

GN FN0291.

OS Fusobacterium nucleatum (subsp. nucleatum).

OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;

OC Fusobacterium.

OX NCBI_TaxID=76856;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 25586;

RX MEDLINE=21886394; PubMed=11889109;

RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,

RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,

RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,

RA Larsen N., D'Souza M., Walunas T., Fusch G., Haselkorn R.,

RA Forstein M., Kyripides N., Overbeek R.;

RT "Genome sequence and analysis of the oral bacterium Fusobacterium

RT nucleatum strain ATCC 25586.";

RL J. Bacteriol. 184:2005-2018(2002).

DR EMEL; AE010541; AAL94497.1; -

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR008638; Haemagg.act.

DR InterPro; IPR005058; Peptidase_S26.

DR InterPro; IPR001452; SH3.

DR Pfam; PF05860; Haemagg.act; 1.

DR Pfam; PF00018; SH3; 1.

DR PROSITE; PS00761; SPASE_I_3; 1.

KW Complete proteome.

SQ SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;

Query Match 6.0%; Score 218.5; DB 16; Length 1881;

Best Local Similarity 21.4%; Pred. No. 0.046;

Matches 175; Conservative 130; Mismatches 306; Indels 205; Gaps 37;

Qy 30 QAPMVVTS-----STGT-----DLSIPSELE---NIPSENOY-FQSAIWSGFI 69

Db 246 QAPVATKGDVVVISGKGVLYLKDQAKRDIKISSTETIGSKLLAENAINKSGKTSNG 305

Qy 70 KVKSDEYTFATSDNHNVTMVDQEVN-----KASNSKIRLEKGR 112

Db 306 QIRANNITNGVDSSNLTITNKDITISGLNKNVSSNLSNKNVVEE-K 364

Qy 113 LYQIKI-----QYQRENPTKGLDFKLYWT--DSQNKKEVSSDNLQLPELKQSS 161

Db 365 LSSTKITNLGLSAKEIEKTNIFNSGKLFKNITAKDFKNNGEV-SSENLTITNLENSN 423

Qy 162 -----NSRKKRSTAGTVPDRNDG----- 182

Db 424 INVKNINSGIVNKTNAEITSKNLTNNLDRNGITIIINVSSGVIANNGKLLVGNITN 483

Qy 183 ---IPDSLEVGGYTDVKNKRTF.LSPWISNTHEKGLTKYKSSPEKRWSTADPYSDPEKV 239

Db 484 SQNLNTATVQKTLIDKNKNSGKILSDNLAKDIFSSGNISAKVITTTQELINSCEII 543

Qy 240 TGRIDKNVSPARHPLVAAVPIVHVMEN-----IILSKNEOSTONTDSETRTSKNT 293

Db 544 SNNLSNNINNSKNIFVNGNLKISNNLANSVGIEGLELNTSIENTGNITIKNKLTSQNL 603

Qy 294 STSRHTSNSSSTVAIDHSLSLAGERTWAETMGINTADTARLANANTRYVNTGTAPIYV 353

Db 604 N-NKNTANVAVAGFLDVHKNISVVG-NIKAITMTKNLND-----NSG-----NI 645

Qy 354 LPTTSLVLGKQTLATIKADENQLSQILAPNYPYPSKNLPIALN-----AQKDASS 405

Db 646 L-TNSLTTAENINKGSIITA-XNISQNLVNSGVSISDNIT-VADNITNINIFANEKISA 702

Qy 406 TPTVNNYQFLEXTKQLRLDTQVYGNIAFYAFENGVRVDTGSNNSEV-----L 457

Db 703 DKIS-NSNK-LVAKNTEITKLTND---GNIVVK--ENLKAKDITNSNTIKVGENLNTDKL 755

Qy 458 POIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLTKALKIATFGNENGLQ 517

Db 756 QNSKTLIAKNINIEKSLANNINGKITSLNANITSDIKNNNGIIQAIK-----NIN 805

Qy 518 YQKDIITBPDFNFDOOTSQNIKNQALAEALNATNIYTVLD---KILNAKONILIRDKRPHY 574

Db 806 IKTSNDLKLKGKYTANDSLINAKSLENNGN-----LENDGKIKFNLGTNLNNKISS 860

Qy 575 DRNNIAY-----GADESIVVKEAHREVINS--TEGLLL-----NI-DKDIR 612

Db 861 SNLNTANEISNNGVNSIGSEANLITANSKLNEGILLFGEGIENKLTGTNTITNGVI 920

Qy 613 KILSGVIVEIETEGLEKVINDRYDMLNISLIRODGKTF-----IDFKK--VYNDKLPY 664

Db 921 SLSGLKLTAKDVNDKHIISDNDLITDVNSITNKGILYSTNNMKVDFKFNLDKAEIY 980

Qy 665 ISNPYKNVYAVTKENTIIINPSENGDSTNGIKKI 700

Db 981 SSG-----DITIN-SENG-TFTNRVGD 1001

RESULT 11

Q897H6 PRELIMINARY; PRT; 675 AA.

ID Q897H6

AC Q897H6

DT 01-JUN-2003 (Tremblrel. 24, Created)

DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)

DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)

DE Membrane-associated protein.

GN CTC00759.

OS Clostridium tetani.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1513;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=Massachusetts / E88;

RX MEDLINE=22457253; PubMed=12552129;

RA Brueggemann H., Baumer S., Fricke W.F., Wierze A., Liesegang H.,

RA Decker I., Herzberg C., Martinez-Arias R., Merki R., Henne A.,

RA Gottschalk G.;

RT "The genome sequence of Clostridium tetani, the causative agent of

RT tetanus disease";

RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).

DR EMBL; AE015938; AAC03560.1; -.

KW Complete proteome.

SQ SEQUENCE 675 AA; 77585 MW; 5572933FDE209BFB CRC64;

Query Match 5.6%; Score 205.5; DB 16; Length 675;

Best Local Similarity 20.5%; Pred. No. 0.051;

Matches 144; Conservative 106; Mismatches 217; Indels 235; Gaps 34;

Qy 42 DLSIPSELENTPSE-NOYFQSAIWSGFIKVK--SDEYTFATSDNHNVTMVDQEVIN 98

Db 52 DINIESK--ENTEGKLCIYISKEKEHHIKISSGVSNKYTFIIVPVKNGV-----KEMKI 103

Qy 99 KASNSNKILEKGRLYQIKIYQRENPTKGLDFKLYWTDSQNKKEVSSDNLQLPELKQ 158

RP	SEQUENCE FROM N.A.	
RC	STRAIN=Serovar 3; PubMed=11048724;	
EX	MEDLINE=20500219; Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,	
RA	Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,	
RA	Cassell G.H.;	
RT	"The complete sequence of the mucosal pathogen Ureaplasma	
RT	urealyticum.";	
RL	Nature 407:757-762(2000).	
DR	EMBL; AE002145; AAF30894.1; -	
KW	Hypothetical protein; Complete proteome.	
SQ	SEQUENCE 4688 AA; 534880 MW; 553ABFAFFEE1997E CRC64;	
	Query Match 5.5%; Score 200; DB 16; Length 4688;	
	Best Local Similarity 20.3%; Pred. No. 1.1;	
	Matches 166; Conservative 127; Mismatches 313; Indels 210; Gaps 37;	
Qy	19 LLGYFSLNFAQAPVMTSTTGDLSPSSLENIIPSENQYFQSAIWSGRIKVKSDYDT 78	
Db	3699 LVDVYLD-NIHQIDEIRKIFKHNH-SKEIENPGVTMISKGNKSPDTTANPEFK 3756	
Qy	79 FATSADNHVTWVDDQEVINKASNKIRLEKRLYGIKIQYQENRTEKGLDFKLYWD 138	
Db	3757 IETQ-----DNDVLNIDATVFKDEHNKIKIVRIKEN-----ND 3795	
Qy	139 SONKEVSSNLQLPKQKSSN-----SRKRSTSGPTVPDRD----- 179	
Db	3796 WLKGGQI---DNLN-PETKYLENIELSKPLKTHNLNSINDKENISLITETGNPVLKV 3851	
Qy	180 ---NDGIPDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSPKSWSTASDPYSD 235	
Db	3852 IOTQNDTINDTQITINVLGVSNSK-YNGROIKVYKDNVNIYESS---LITLQKQND 3907	
Qy	236 FEKVTGRIDKNVSPARHPLVAAPVHVDMENILSKN-EDQSTQNTDSET-RTISKNT 293	
Db	3908 YQLL-SNLNSN-----REYFEKIEINHSINTNPFDEKLKGVSNFTITQTKNT 3957	
Qy	294 STSRHTSNSNSTVAIDHLSLAG----- 318	
Db	3958 TVQWDSATVTRGVNFNFKIKSEDKILENNQOVVAPKPTIERTNTWLOYTRPLK 4017	
Qy	319 -----BRTWAETMGLNT---ADTARLANIRYVNTGTAPVYNVLTSLVLKQNTLA 368	
Db	4018 DVTDFEKGFTWAHDLNSVNFKEETTYKLVKIQFVNKPTAKNNINNSNNVILDNTNSI 4077	
Qy	369 -----TIKADENQLSOLAPNNYPSKVLAPALN-AQKASSTPITMAY--NQPLELE 419	
Db	4078 NSNYEFTTKVGHKLINITSNNVNTNSQTINFTLGVKKSWSVGKKIKLSYKSNDSSES 4137	
Qy	420 KTKQLRLTDQVYGNIAIYNFENGK-----VRVDTGSNWSSEVLPIQIETTARIIFNGKD 473	
Db	4138 HTNEVLIESNKTQYNILLNNLKRNTYTLIDVKLIDNNVSDFPKEGNLITSFITTSTA 4197	
Qy	474 LNLVERRAAVNPSPLETT-----KPDMLKEA-----LKIAGFNEPNEGLO 517	
Db	4198 INVNLIEISNRASTNLKSTIKINLNDPDVLRDKQOATIVYGNKQAMGFTVSGNIK 4257	
Qy	518 YQCKDITDFDNFDQOOTSONIK-NQLAELNATNIYVLDTKIKLAKQNNILRDKRFYVD- 575	
Db	4258 YLTALVDLNFN-DKVNIVNISPNNKPSIAEN-----IGDKSNII-----YNDSD 4304	
Qy	576 -----RNNIAGA---DESVVKEAHREVINSSTEGLLNLTDKDIRKILSYGIVIEIDTE 626	
Db	4305 IPKLEINNDIIVNGPINEIIVYKANQK--NNIDVDLGLQINPKIAHLNR-FIAKFKSTN 4361	
Qy	627 GLKEVIND--RYDMLNLSL-RQDCKTRIDF-----KKY----- 657	
Db	4362 -----NDLIETVINGSLSLVNDGKTSIRFTLNLKANKULYSLVVYLVNNSNTIVE 4415	
Qy	658 NDKFLYISNPNKYVAVTKENTIIINPSENGDTS 693	
Db	4416 SNKLP-KLNNINYQIK--INKSHITII--SKNGEWS 4445	

RESULT 14	
O8IHY4	PRELIMINARY; PRT; 2849 AA.
ID	Q8IHY4;
AC	Q8IHY4;
DT	01-MAR-2003 (TREMELrel. 23, Created)
DT	01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT	01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE	Hypothetical protein.
GN	PF11_0392.
OS	Plasmodium falciparum (isolate 3D7).
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=36329;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=3D7;
RC	MEDLINE=22255705; PubMed=12368864;
RA	Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA	Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA	Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA	Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA	Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA	Marletta D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA	McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA	Fraser C.M., Barrell B.;
RT	"Genome sequence of the human malaria parasite Plasmodium
RT	falciparum.";
RL	Nature 419:498-511(2002).
DR	EMBL; AE014841; AAN35975.1; -
KW	Hypothetical protein.
SQ	SEQUENCE 2849 AA; 335916 MW; B5515D173D96B813 CRC64;
	Query Match 5.5%; Score 199.5; DB 5; Length 2849;
	Best Local Similarity 20.6%; Pred. No. 0.63;
	Matches 169; Conservative 128; Mismatches 322; Indels 201; Gaps 39;
Qy	3 KQENLLNSESSESSGGLGY-YFSDLNQ--APVMTSTTGDLSP--SSLENIIPSEN 57
Db	1862 KHNVTFFKEENLKSIFKYNNNNNINEKDDIPKSIQSSFINEDNIATYKNEGNMUKLN 1921
Qy	58 Q-YFQSAIWSGFIKVKSDYTFATSDNHVTMWYDD-----QEVINKASNSK---- 105
Db	1922 EDIIQERFHHNIKTVMNKNCSYSDCNIVKINYDELNDSTQTKELNEGKSNKGAEAW 1981
Qy	106 -IIELEKRLY-QIKIQYQENRTEKGLDFKLYWTSQNKKEVISSDNLQLPKQKSSNS 163
Db	1982 IIDIKNETHYPIKIDKKEKNEDEK--RNKYMYNNKNDNKNIMKGSN---KSMKNNKNS 2036
Qy	164 RKKRSTSGPTVPDRDNDGIPDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSSP 223
Db	2037 NKMKH-----IPLSVNKKY--NKSSINKKYENNIN-----KXNK 2069
Qy	224 EKWSTASDPYSDF--EKVTGRI-DKNVSPARHPLVAAPVHVDMENILSKNEDQSTQ 280
Db	2070 DKLNLVNSISKLVSQKIKQELSNKNSK-----DILNFEITKIKKSKK 2114
Qy	281 NTDSERTISKNTSRTSTNS-----NSSTVA-----IDHLSL 316
Db	2115 ETKNTHNTKNN 316
Qy	317 AGE---RTWAETMGLNTA---DTARLANA-----NIRVNTGTAPIYVNLPTTSIV 360
Db	2175 APDGIYTWVNNIDTVMYKRYFDSKKNINHVPLNVPYLN-----NDLLTNVYI 2227
Qy	361 LGKQNTLATIKADENQLSQ-----ILAPNNYPSKVLAPALNAQKASSTPITMAYNQFL 416
Db	2228 LQNN 2277
Qy	417 ELEKTKQLRLDT-----DQVYGNIAIYNFENGKRVVDT-----GSNWSSEVL 458
Db	2278 NMQLLNHNLNNSINVIDQNYLSYNNLNCINGNINKDCTKDIVIGIPNNNTNQOI--P 2336
Qy	459 QIQTETARIIFNGKOLNVLERRIAAVNPSPDPLE-----TTKPDMLKEALKIAFGF- 509

```
Db 2337 TI-ELDDTILKN--DUNLILNNVNVHSVVEMLNNTQNVQKLYNDIQENLHITNSLH 2393
QY 510 -NEPNGNIQYQOKD-----ITFDNFDDQOTSQNIKNQLAELNATNIYTVLDKIKL 559
Db 2394 NNNNNNNMGLGCVLNSYLFNINSFNPMMAYIYNNNNNNNNNNNNNNNNNNNNNNNN 2453
QY 560 NAK-----MNLIRDKRHYDRNNIAGVADSVVKEAHREVINSSTEGLLANIDKIRKI 614
Db 2454 NVKNEIINNVSIEDTNEKNEG-----MDKNKYIINK--ENNKNNNNNIDII 2502
QY 615 LSGYIVEIEDTEGLKEVINDRYDMLNISLRQDGKTFIDFKYNDKPLPLYISNPNYKVV 674
Db 2503 SNGQNNNTNANTMEKNNNRDDDIYKQNMKRSNLSDFKKLDSE-----KKKSNWI 2556
QY 675 ---YAVTKENT--IINPENGDTSTNGIKKILIFSKGYE 709
Db 2557 BEGARKMTNTQLLNYSN---RKKGLRDMSTYADKVL 2593

RESULT 15
Q81104
ID Q81104 PRELIMINARY; PRT; 3468 AA.
AC Q81104;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pf11_0371.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Hatt D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RA "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511 (2002).
DR EMBL; AE014841; AAN35955.1; -.
KW Hypothetical protein.
SQ SEQUENCE 3468 AA; 411913 MW; 506F7D62999BA7B1 CRC64;
```

```
Query Match 5.5%; Score 199.5; DB 5; Length 3468;
Best Local Similarity 19.0%; Pred. No. 0.81;
Matches 157; Conservative 130; Mismatches 263; Indels 275; Gaps 39;

QY 55 SENQYQSAIWSGFIKVKKSDEYTFATSDNHTVMWDDQEV---INKASNSNKIRLEKG 111
Db 160 SENY-----TKNDDKYNNSTNSGY-----NKEIEFLKNIAKSHSLKKIEKN 202
QY 112 RLYQIK-----IOYQRENTEKGLDPK---LYW---TDSQ-----KKEVISSNVL 151
Db 203 INLLKXCDSDNNDLNLEFKNNVHLNNDYRSEDLFHDNRVNNQNNLKDIIDYIHDNI 262
QY 152 QLPKQKSSNRKSKKSTESAGPTVPDRDNDGIPDSLEVEGYVDV-KNKRFTLSPWISNI 210
Db 263 MI-----ESNNDRNDKCNFKNTNPNYERHIIIVDTKNKGNNNNNNSNPFDS--ISNI 315
QY 211 HEKKGITKYSGPEKWSIASDP-----YSDFE-KVTGRIDKNVSPPEARHPLVAAYPIVHV- 264
Db 316 NKK-----ISYPINMYNSSEDKTNLQINLMDLSILSDSLKNAYSFYSLIS 360
QY 265 ---DMENIILSKNQDQSTQNTD--SETRTISKNTSRTHTS-----NSNS 305
```

```
Db 361 EKNSNNGSHKNNMYRVDIVIDINISKFSNISLHDSMDNNEFNSLCSFNSS 420
QY 306 STVAIDHSLSLAGERTWAETMGL-----NTADTARLANANTRY-----VNTGTAPIYV 353
Db 421 KCVKDEMITQFVGNRKYVKTMDANDNYMKNSIEQHSINMESNHFKNQKNGSKEDAND 480
QY 354 LPTTSVLGKNQ-----TLATIKADE--NOLSOILAPNN----- 385
Db 481 NLNSSHILNNNGIGQVNSNLSLDYLSIKQSSMNLNNEALNNINVTNNNGSSNNKH 540
QY 386 -----YYPSSKNLAPIALNAQKADASTPTITMNYNQ---FLELEKTKOL--RLDTP 429
Db 541 NSNVKTSQVYVNPNDNIN--NMHQNLNLSYMKNSNLTNSGFKKIPKNKNIISNIDPD 598
QY 430 Q-----VYQNI-----ATYNPENGVRVDTGSNMSEVLPO-QUETAR----- 466
Db 599 NNIFKSYIKENVVKNOESNINHQVHEKNTYNDIEIKN-----NNIENNTQNTCNFIN 554
QY 467 --IIFNGKDLNLVERRIAAVNPSPDPLETTK-----PDMTLKEA 502
Db 655 NDDIINKKKIKETIKYKIDSISLLNDLSLNKLESNSSIMDRYTKNNYEKFLDDVILDDS 714
QY 503 L-----KIAFGFNPNGNLOYOGKOITEFDNFDDQOTSQNIKNQLAEL 545
Db 715 IFATSNELLQHSNYTTTNNHIFDNNNNNNNNINQEKOLFONDYN---KETYNNNIMLSN 771
QY 546 NATNIYTVLDKIKLAKNNILIRDKFHYDRNNIAGVADSVVKEAHREVINSSTEGLL 605
Db 772 NAEMLF-----KISYSCNDLVLGKNEIILD--RNVENSKTE--- 806
QY 606 NIDKDIRKILSGYIVEIEDTEGLKE-----VINDRYDMLNISS---LRDGKTFIDF 654
Db 807 -----QYEEYQNKEDIILKYHKDNIIDNNYNDVNIKKDCHLKMNDQDNIKI 854
QY 655 KKYND---KLPL-----YISNPNYKVN-VYAVTKENTIIINPSEN 689
Db 855 NKSQDSKKNKPINNNECNVHNKQFKINEIDHFKEENTLNYESKN 899

Search completed: May 3, 2004, 19:40:09
Job time : 36.4955 secs
```

```

QY 657 NRYDWMNLSSLRQDGKTFIDFKYNDKLPYISNENYKVVAVTKENTIIINPSENGDT 716
DB 641 NRYDWMNLSSLRQDGKTFIDFKYNDKLPYISNENYKVVAVTKENTIIINPSENGDT 700
QY 717 STNGIKKILIFSKGYEIG 735
DB 701 STNGIKKILIFSKGYEIG 719

RESULT 8
PCT-US94-01624-31
; Sequence 31, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; STREET: Plaza
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01624-31

Query Match 94.8%; Score 3576; DB 5; Length 719;
Best Local Similarity 95.1%; Pred. No. 3.8e-249;
Matches 703; Conservative 4; Mismatches 8; Indels 24; Gaps 2

QY 1 EVKQENLLNSESSESSGCLGYFSDNLNFQAPMVVTSSTTGDLSIPSELENIIPSENYQF 60
DB 1 EVKQENLLNSESSESSGCLGYFSDNLNFQAPMVVTSSTTGDLSIPSELENIIPSENYQF 60
QY 61 QSAIWSGFIKVKKDEYTFATSDANHHVTMWDDDEVINKASNSNKIRLEKGLRYQIKIY 120
DB 61 QSAIWSGFIKVKKDEYTFATSDANHHVTMWDDDEVINKASNSNKIRLEKGLRYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVIGSDNIQLPELKQKSGNS----RKKSTASGPTVP 176
DB 121 QRENPTKGLDFKLYWTDSONKKEVIGSDNIQLPELKQKSGNSNTATIMNQRGNFLQGPTVP 180
QY 177 DRDNDGIPDSLEVGSGYVDVKNKKTFLSPWLSNIHEKGLTKYKSSPEKWSSTASDPYSDF 236
DB 181 DRDNDGIPDSLEVGSGYVDVKNKKTFLSPWLSNIHEKGLTKYKSSPEKWSSTASDPYSDF 240

```

Db	421	LNAQDFFSTPTIMYNOFLELEKTKQLRLDQVYGNATYNFENGKRVVDTGSNWSEV	480
Qy	481	LPOIQTETARIIFNGKDLNVLERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL	540
Db	481	LPOIQTETARIIFNGKDLNVLERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL	540
Qy	541	QYQGXDIETEFNFPQQTSONIKQLAELNATNIYTVLDKTKLNAXNNILIRDKRFHYDR	600
Db	541	QYQGXDIETEFNFPQQTSONIKQLAELNATNIYTVLDKTKLNAXNNILIRDKRFHYDR	600
Qy	601	NNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY	660
Db	601	NNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY	660
Qy	661	DMLNLSLRQDGKTFIDFKKXNDKPLYSINPNYKVNVIYAVTKNTIINPSENGDTSTNG	720
Db	661	DMLNLSLRQDGKTFIDFKKXNDKPLYSINPNYKVNVIYAVTKNTIINPSENGDTSTNG	720
Qy	721	IKKIL---IFSKKG 731	
Db	721	IKKILKVLGKKG 734	
RESULT 6			
PCT-US94-01624-12			
; Sequence 12, Application PC/TUS9401624			
; GENERAL INFORMATION:			
; APPLICANT: Leppla, Stephen H.			
; APPLICANT: Klimpel, Kurt R.			
; APPLICANT: Arora, Naveen			
; APPLICANT: Singh, Yogendra			
; APPLICANT: Nichols, Peter J.			
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND			
; TITLE OF INVENTION: RELATED METHODS			
; NUMBER OF SEQUENCES: 31			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW			
; STREET: Stewart Street Tower, 20th Floor, One Market			
; CITY: San Francisco			
; STATE: CA			
; COUNTRY: USA			
; ZIP: 94105			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; OPERATING SYSTEM: IBM PC compatible			
; SOFTWARE: Patent in Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: PCT/US94/01624			
; FILING DATE: June 25, 1993			
; CLASSIFICATION:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Weber, Kenneth A.			
; REGISTRATION NUMBER: 31,677			
; REFERENCE/DOCKET NUMBER: 15280-115			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (415) 543-9600			
; TELEFAX: (415) 543-5043			
; INFORMATION FOR SEQ ID NO: 12:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 903 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
PCT-US94-01624-12			
Query Match 98.3%; Score 3708.5; DB 5; Length 903;			
Best Local Similarity 98.8%; Pred. No. 1.5e-258;			
Matches 725; Conservative 1; Mismatches 5; Indels 3; Gaps 1;			
Qy	1	EVKQENRLNSESQGLLYGYFSDLNQAPMVVTSSTTGLSIPSSSELENIPSENQYF	60

```
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 903 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-021-601-12

Query Match          98.3%; Score 3708.5; DB 1; Length 903;
Best Local Similarity 98.8%; Pred. No. 1.5e-258;
Matches 725; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKYKSDRYTATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKYKSDRYTATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTASGTPVDRDN 180
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWTASDPYDFEKT 240
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVLPVTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYVLPVTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKDASSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIATFGNPNGL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIATFGNPNGL 540
QY 541 QYQKDIPTDFNFDQOTSQNIKNQLAELNATNIYTVLDKIKLNKQNLILIRDKRFHYDR 600
Db 541 QYQKDIPTDFNFDQOTSQNIKNQLAELNATNIYTVLDKIKLNKQNLILIRDKRFHYDR 600
QY 601 NNTAVGADSVKAEHREVINSSTEGLLNIDKDIRKILSGYVEIETEGELKEVINDRY 660
Db 601 NNTAVGADSVKAEHREVINSSTEGLLNIDKDIRKILSGYVEIETEGELKEVINDRY 660
QY 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNKYVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNKYVAVTKENTIIINPSNGDTSTNG 720
QY 721 IKKIL----IFSKKG 731
Db 721 IKKILKKVILGKKG 734

RESULT 5
US-08-082-849B-12
; Sequence 12, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
```

```
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,849B
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,601
; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-161-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 903 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-082-849B-12

Query Match          98.3%; Score 3708.5; DB 1; Length 903;
Best Local Similarity 98.8%; Pred. No. 1.5e-258;
Matches 725; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKYKSDRYTATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKYKSDRYTATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTASGTPVDRDN 180
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWTASDPYDFEKT 240
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVLPVTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYVLPVTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKDASSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
```


QY 661 DMLNSSLRQDGKTFIDFKKNDKLPYISNPNYKVNVAVTKENTIIINPSENGDSTNG 720
DB 661 DMLNSSLRQDGKTFIDFKKNDKLPYISNPNYKVNVAVTKENTIIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 3
PCT-US94-01624-4
; Sequence 4, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW
; STREET: Stewart Street Tower, 20th Floor, One Market
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01624-4

Query Match 99.5%; Score 3753; DB 5; Length 735;
Best Local Similarity 99.6%; Pred No. 7, 2e-262;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLPNFAPMVTSSTGDLSPSELENIPSENYF 60
DB 1 EVKQENRLNSESQGLGYFSDLPNFAPMVTSSTGDLSPSELENIPSENYF 60
QY 61 QSAINSGFTKVKSDBYTATADNHVTMWVDDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAINSGFTKVKSDBYTATADNHVTMWVDDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTPVDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTPVDRDN 180
QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEY 240
DB 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEY 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTATARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTATARL 360
QY 361 NAMIRYNTGTAPIYNNVLTPTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NAMIRYNTGTAPIYNNVLTPTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTITMNNYNOLELEKTKQLRLDTPQVYGNATYNFENGRVVDGSKWSEV 480
DB 421 LNAQDDFSSPTITMNNYNOLELEKTKQLRLDTPQVYGNATYNFENGRVVDGSKWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAFGNEPENGNL 540
DB 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAFGNEPENGNL 540
QY 541 QYQKDIETDFNFDQOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNLIIRDKRPHYDR 600
DB 541 QYQKDIETDFNFDQOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNLIIRDKRPHYDR 600
QY 601 NNTAVGADESVMKEAHREVINSSTGLLNIDKIRKILSGYIYEIETDEGLKEVINDRY 660
DB 601 NNTAVGADESVMKEAHREVINSSTGLLNIDKIRKILSGYIYEIETDEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKNDKLPYISNPNYKVNVAVTKENTIIINPSENGDSTNG 720
DB 661 DMLNSSLRQDGKTFIDFKKNDKLPYISNPNYKVNVAVTKENTIIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 4
US-08-021-601-12
; Sequence 12, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 133 Carnegie Way, Suite 400
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,601
; FILING DATE: 19930212
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880

Db 1 EVKQENRLNSESQGLLYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDPLKLYWDSQNKKEVSSDNLOLPELKOKGSNSRKRSTSGPTVPDRDN 180
Db 121 QRENTEKGLDPLKLYWDSQNKKEVSSDNLOLPELKOKGSNSRKRSTSGPTVPDRDN 180
QY 181 DGIPOSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASPYSDFEKVT 240
Db 181 DGIPOSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASPYSDFEKVT 240
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDDASSTPTIMYNNQFLEKTKQLRLDQVYGNIAATYNFENGRVVDTSNWSSEV 480
Db 421 LNAQKDDASSTPTIMYNNQFLEKTKQLRLDQVYGNIAATYNFENGRVVDTSNWSSEV 480
QY 481 LPQIQETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLTKEALKIAFGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLTKEALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFDOOTSQNKQLAELNATNITVLDKIKLAKXNILLIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNKQLAELNATNITVLDKIKLAKXNILLIRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAREHREINVESTGLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAREHREINVESTGLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNITSLQDQKTFIDFKYNDKPLYSNPNYKNTYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNITSLQDQKTFIDFKYNDKPLYSNPNYKNTYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 2

US-08-082-849B-4
; Sequence 4, Application US/08082849B
; Patent No. 5677274

GENERAL INFORMATION:

; APPLICANT: Leppa, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Arcora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,849B
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,601
; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-161-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-082-849B-4

Query Match 99.5%; Score 3753; DB 1; Length 735;

Best Local Similarity 99.6%; Pred. No. 7.2e-262; Indels 0; Gaps 0;
Matches 732; Conservative 0; Mismatches 3;

QY 1 EVKQENRLNSESQGLLYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLLYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDPLKLYWDSQNKKEVSSDNLOLPELKOKGSNSRKRSTSGPTVPDRDN 180
Db 121 QRENTEKGLDPLKLYWDSQNKKEVSSDNLOLPELKOKGSNSRKRSTSGPTVPDRDN 180
QY 181 DGIPOSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASPYSDFEKVT 240
Db 181 DGIPOSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASPYSDFEKVT 240
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDDASSTPTIMYNNQFLEKTKQLRLDQVYGNIAATYNFENGRVVDTSNWSSEV 480
Db 421 LNAQKDDASSTPTIMYNNQFLEKTKQLRLDQVYGNIAATYNFENGRVVDTSNWSSEV 480
QY 481 LPQIQETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLTKEALKIAFGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLTKEALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFDOOTSQNKQLAELNATNITVLDKIKLAKXNILLIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNKQLAELNATNITVLDKIKLAKXNILLIRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAREHREINVESTGLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAREHREINVESTGLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:31:43 ; Search time 15.0986 Seconds
(without alignments)
2513.152 Million cell updates/sec

Title: US-09-848-909A-13

Perfect score: 3772

Sequence: 1 EVKQENRLNSESQGLL.....FTNGIKKILFFSKKGVEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/aaa/5A COMB.pep:*

2: /cgn2_6/prodata/2/aaa/5B COMB.pep:*

3: /cgn2_6/prodata/2/aaa/6A COMB.pep:*

4: /cgn2_6/prodata/2/aaa/6B COMB.pep:*

5: /cgn2_6/prodata/2/aaa/PCFUS COMB.pep:*

6: /cgn2_6/prodata/2/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3753	99.5	735	1	US-08-021-601-4
2	3753	99.5	735	1	US-08-082-849B-4
3	3753	99.5	735	5	PCT-US94-01624-4
4	3708.5	98.3	903	1	US-08-021-601-12
5	3708.5	98.3	903	1	US-08-082-849B-12
6	3708.5	98.3	903	5	PCT-US94-01624-12
7	3576	94.8	719	1	US-08-082-849B-31
8	3576	94.8	719	5	PCT-US94-01624-31
9	1275	33.8	288	4	US-09-273-839A-8
10	767	20.3	881	3	US-08-960-780-32
11	767	20.3	881	3	US-09-073-898-32
12	767	20.3	881	4	US-09-307-106-8
13	767	20.3	881	4	US-09-850-351A-32
14	765	20.3	884	1	US-08-471-033-5
15	765	20.3	884	2	US-08-471-044-5
16	765	20.3	884	2	US-08-463-483A-5
17	765	20.3	884	2	US-08-471-046A-5
18	765	20.3	884	2	US-08-470-566B-5
19	765	20.3	884	2	US-08-469-334-5
20	765	20.3	884	2	US-09-300-529-5
21	765	20.3	1346	1	US-08-471-033-23
22	765	20.3	1346	2	US-08-471-044-23
23	765	20.3	1346	2	US-08-463-483A-23
24	765	20.3	1346	2	US-08-471-046A-23
25	765	20.3	1346	2	US-08-470-566B-23
26	765	20.3	1346	2	US-08-469-334-23
27	765	20.3	1346	3	US-09-300-529-23

28	764	20.3	852	1	US-08-471-033-36	Sequence 36, Appl
29	764	20.3	852	2	US-08-471-044-36	Sequence 36, Appl
30	764	20.3	852	2	US-08-463-483A-36	Sequence 36, Appl
31	764	20.3	852	2	US-08-471-046A-36	Sequence 36, Appl
32	764	20.3	852	2	US-08-470-566B-36	Sequence 36, Appl
33	764	20.3	852	2	US-08-469-334-36	Sequence 36, Appl
34	764	20.3	852	3	US-09-300-529-36	Sequence 36, Appl
35	764	20.3	1338	1	US-08-471-033-50	Sequence 50, Appl
36	764	20.3	1338	2	US-08-471-044-50	Sequence 50, Appl
37	764	20.3	1338	2	US-08-463-483A-50	Sequence 50, Appl
38	764	20.3	1338	2	US-08-471-046A-50	Sequence 50, Appl
39	764	20.3	1338	2	US-08-470-566B-50	Sequence 50, Appl
40	764	20.3	1338	2	US-08-469-334-50	Sequence 50, Appl
41	764	20.3	1338	3	US-09-300-529-50	Sequence 50, Appl
42	754.5	20.0	784	3	US-09-371-913A-7	Sequence 7, Appl
43	754.5	20.0	784	4	US-09-967-805-7	Sequence 7, Appl
44	745.5	19.8	860	4	US-08-307-106-48	Sequence 48, Appl
45	735.5	19.5	834	1	US-08-471-033-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-08-021-601-4

; Sequence 4, Application US/08021601

; Patent No. 5591631

; GENERAL INFORMATION:

; APPLICANT: Leppla, Stephen H.

; APPLICANT: Klimpel, Kurt R.

; APPLICANT: Nichols, Peter J.

; APPLICANT: Arora, Naveen

; APPLICANT: Singh, Yogendra

; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND

; TITLE OF INVENTION: RELATED METHODS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Needle & Rosenberg, P.C.

; STREET: 133 Carnegie Way, Suite 400

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: USA

; ZIP: 30303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/021,601

; FILING DATE: 19930212

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Spratt, Gwendolyn D.

; REGISTRATION NUMBER: 36,016

; REFERENCE/DOCKET NUMBER: 1414.057

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 404/688-0770

; TELEFAX: 404/688-9880

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 735 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-021-601-4

Query Match 99.5%; Score 3753; DB 1; Length 735;

Best Local Similarity 99.6%; Pred. No. 7.2e-262;

Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYFSDLNFOAPVWVTSSTTGGDLSIPSELENIPSENOYF 60

QY 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKEHTIINPSENGDTSTNG 720
DB 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKEHTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 15
AAV56959
ID AAY56959 standard; protein; 736 AA.
XX
AC AAY56959;
XX
DT 25-APR-2000 (first entry)
XX
DE B. anthracis MAT-PA protein.
XX
KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;
KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.
XX
OS Bacillus anthracis.
XX
PN WO200002522-A2.
XX
PD 20-JAN-2000.
XX
PF 09-JUL-1999; 99WO-US015568.
XX
PR 10-JUL-1998; 98US-0092416P.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;
XX WPI; 2000-182165/16.
DR N-PSDB; AAZ56875.
XX
PT Recombinant DNA construct useful as vaccines for anthrax, in producing
PT host cells for analyzing the drugs and agents inhibiting anthrax.
XX
PS Disclosure; Page 34; 35pp; English.
XX
CC The invention provides a recombinant DNA construct that comprises a
CC vector and at least one nucleic acid (or its fragment) encoding a
CC combination of Bacillus anthracis proteins, selected from protective
CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA
CC with its secretory signals replaced with those of tissue plasminogen
CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine
CC for anthrax and in producing infectious alpha virus particles. These
CC particles, expressing the B. anthracis proteins are useful also as
CC vaccines for anthrax. Host cells transformed with the construct are
CC useful for analyzing the effectiveness of drugs and agents that inhibit
CC anthrax or B. anthracis proteins. The present sequence represents a B.
CC anthracis MAT-PA protein
SQ Sequence 736 AA;

Query Match 99.5%; Score 3753; DB 3; Length 736;
Best Local Similarity 99.6%; Pred. No. 1.6e-247;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLLYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
DB 2 EVKQENRLNSESQGLLYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENQYF 61
QY 61 QSATWSGPIKVKSDYEYTFATSDNHVTWDDQEVINKASNSNKIRLEKGRLYQIKIY 120
DB 62 QSATWSGPIKVKSDYEYTFATSDNHVTWDDQEVINKASNSNKIRLEKGRLYQIKIY 121
QY 121 QRENPTKGLDFKLYWTDSONKKEVI6SDNLIQLPELKQKSSNRKKGTSAGPTVPDRN 180

DB 122 QRENPTKGLDFKLYWTDSONKKEVI6SDNLIQLPELKQKSSNRKKGTSAGPTVPDRN 181
QY 181 DGI PDSLEVEGYTVDVKNETFTFLSPWISNIHEKGLTKYKSSPEKWTASDYPSEKVT 240
DB 182 DGI PDSLEVEGYTVDVKNETFTFLSPWISNIHEKGLTKYKSSPEKWTASDYPSEKVT 241
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 242 GRIDKNVSPARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 301
QY 301 SEVHGNAEVHASFDDIGGSVSAFSSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 302 SEVHGNAEVHASFDDIGGSVSAFSSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 361
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPPSKNLAPIA 420
DB 362 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPPSKNLAPIA 421
QY 421 LNAQKDASSPTITWYNOFLEKTKQLRLDTPQVYGNATYNFENGVRVVDTSNWSSEV 480
DB 422 LNAQKDASSPTITWYNOFLEKTKQLRLDTPQVYGNATYNFENGVRVVDTSNWSSEV 481
QY 481 LPOIQTETARIIFNGKOLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB 482 LPOIQTETARIIFNGKOLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAGFNEPENG 541
QY 541 QYQKDIETEDFNFDOOTSONIKNQLAELNATNITVLDKIKLNAKONILIRKRPHYDR 600
DB 542 QYQKDIETEDFNFDOOTSONIKNQLAELNATNITVLDKIKLNAKONILIRKRPHYDR 601
QY 601 NNIAGADESVVKAHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 602 NNIAGADESVVKAHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 661
QY 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKEHTIINPSENGDTSTNG 720
DB 662 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKEHTIINPSENGDTSTNG 721
QY 721 IKKILIFSKKGYEIG 735
DB 722 IKKILIFSKKGYEIG 736

Search completed: May 3, 2004, 19:36:03
Job time : 52.1208 secs

Db 61 QSAIWSGFIKVKSSDEYFATSDAHVYTWVDDQEVINKASNSKIRLEKGLVQIKQY 120
QY 121 QRENTEKGLDFKLWYTSQNKKEVSSDNLOLPELKOKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENTEKGLDFKLWYTSQNKKEVSSDNLOLPELKOKSSNSRKRSTSGPTVPDRDN 180
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSASPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDSASPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
QY 541 QYQKDIITEFDNFDOQTQSNIKQLAELNATNIYTVLDKIKLNKXNILLIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTQSNIKQLAELNATNIYTVLDKIKLNKXNILLIRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGTDFDKKNDKPLIYSNPNYKVNVAVTKEINTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGTDFDKKNDKPLIYSNPNYKVNVAVTKEINTIINPSENGDTSTNG 720
QY 721 IKKILIPSKKGYEIG 735
Db 721 IKKILIPSKKGYEIG 735

RESULT 14
AAM51499
ID AAM51499 standard; protein; 735 AA.
XX AC AAM51499;
XX AC AAM51499;
DT 01-FEB-2002 (first entry)
XX DE Anthrax PA F427D mutant.
XX OS Bacillus anthracis.
KW Anthrax; PA: protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.
XX OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 427 /note= "Wild-type Phe substituted by Asp"
XX WO200182788-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US014372.

XX PR 04-MAY-2000; 2000US-0201800P.
XX PA (HARD) HARVARD COLLEGE.
XX PI Collier RJ, Sellman BR;
XX DR WPI; 2002-017725/02.
XX PT Protecting humans against anthrax using mutant B groups (anthrax
XX PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX PT anthracis.
XX PS Claim 4; Page: 77pp; English.
XX CC The invention relates to antibacterial agents comprising mutant forms of
XX CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX CC in the B moiety of the pore-forming binary A-B anthrax toxin where the B
XX CC moiety is anthrax protective antigen (PA) and using these mutants or
XX CC compositions of them for protecting against Bacillus anthracis infections
XX CC in humans, especially as vaccines. Note: The present sequence is not
XX CC given in the specification but is derived from the Bacillus anthracis
XX CC wild-type PA protein sequence shown in figure 13 (AAM51483)
XX SQ Sequence 735 AA;
Query Match 99.5%; Score 3753; DB 5; Length 735;
Best Local Similarity 99.6%; Pred. No. 1.6e-247;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EVKQENLLNESSSQGLLGYFSDLNFOAPMVVTSSTGDLSPSSELENIPSENOYF 60
Db 1 EVKQENLLNESSSQGLLGYFSDLNFOAPMVVTSSTGDLSPSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSSDEYFATSDAHVYTWVDDQEVINKASNSKIRLEKGRUYQIKQY 120
Db 61 QSAIWSGFIKVKSSDEYFATSDAHVYTWVDDQEVINKASNSKIRLEKGRUYQIKQY 120
QY 121 QRENTEKGLDFKLWYTSQNKKEVSSDNLOLPELKOKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENTEKGLDFKLWYTSQNKKEVSSDNLOLPELKOKSSNSRKRSTSGPTVPDRDN 180
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSASPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDSASPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
QY 541 QYQKDIITEFDNFDOQTQSNIKQLAELNATNIYTVLDKIKLNKXNILLIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTQSNIKQLAELNATNIYTVLDKIKLNKXNILLIRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

PA	(HARD) HARVARD COLLEGE.	
XX		
PI	Collier RJ, Sellman BR;	
XX		
XX	WPI; 2002-017725/02.	
DR		
XX		
PT	Protecting humans against anthrax using mutant B groups (anthrax	
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus	
PT	anthracis.	
XX		
PS	Claim 4; Page; 77pp; English.	
XX		
XX	The invention relates to antibacterial agents comprising mutant forms of	
CC	pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants	
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B	
CC	moiety is anthrax protective antigen (PA) and using these mutants or	
CC	compositions of them for protecting against Bacillus anthracis infections	
CC	in humans, especially as vaccines. Note: The present sequence is not	
CC	given in the specification but is derived from the Bacillus anthracis	
CC	wild-type PA protein sequence shown in figure 13 (AAM51483)	
XX		
SQ	Sequence 735 AA;	
	Query Match 99.5%; Score 3753; DB 5; Length 735;	
	Best Local Similarity 99.6%; Pred. No. 1.6e-247;	
	Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	1 EVKQENRLNSESSESSQGLLGYYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60	
Db	1 EVKQENRLNSESSESSQGLLGYYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60	
QY	61 QSAIWSGFTKVKSDYEYFATSADNHVTMWDDQEVINKASNSKIRLEKRLYQIKIY 120	
Db	61 QSAIWSGFTKVKSDYEYFATSADNHVTMWDDQEVINKASNSKIRLEKRLYQIKIY 120	
QY	121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLPELKQKSNRSKRSTSGTVPDRDN 180	
Db	121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLPELKQKSNRSKRSTSGTVPDRDN 180	
QY	181 DGIPTDSEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWASTASDPSDFEKT 240	
Db	181 DGIPTDSEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWASTASDPSDFEKT 240	
QY	241 GRIDKNVSPKARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300	
Db	241 GRIDKNVSPKARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300	
QY	301 SEVHGNAEVHASFDDIGGSVSGFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
Db	301 SEVHGNAEVHASFDDIGGSVSGFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
QY	361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNYYPSKNLAPIA 420	
Db	361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNYYPSKNLAPIA 420	
QY	421 LNAQKDASTPTIMYNNQFLEKTKQLDLTDQVYGNATYFNPGVRVDTGNSWSEV 480	
Db	421 LNAQKDASTPTIMYNNQFLEKTKQLDLTDQVYGNATYFNPGVRVDTGNSWSEV 480	
QY	481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIATFGFNEPGNL 540	
Db	481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIATFGFNEPGNL 540	
QY	541 QYQKDIITBDFNFDQTSQNTKQLAELNATNIYTVLDKIKLNAMNILLRDKGFHYDR 600	
Db	541 QYQKDIITBDFNFDQTSQNTKQLAELNATNIYTVLDKIKLNAMNILLRDKGFHYDR 600	
QY	601 NNTAVGADESVMKEAHREVINSTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660	
Db	601 NNTAVGADESVMKEAHREVINSTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660	
QY	661 DMLNISLRQDGKTFIDFKKYNKLPXY-SNPYKVNVAVTKENTIIINPSNGTSTNG 720	

Db	661 DMLNISLRQDGKTFIDFKKYNKLPXYISNPNYKVNVAVTKENTIIINPSNGTSTNG 720	
QY	721 IKKILIFSKKGYEIG 735	
Db	721 IKKILIFSKKGYEIG 735	
RESULT 13		
AAM51494		
ID	AAM51494 standard; protein; 735 AA.	
XX		
AC	AAM51494;	
XX		
DT	01-FEB-2002 (first entry)	
XX		
DE	Anthrax PA mutant K395D/K397D/b425K/D426K.	
XX		
KW	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;	
XX	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.	
XX	Bacillus anthracis.	
OS	Synthetic.	
XX		
XX	Key Location/Qualifiers	
FT	Misc-difference 395 /note= "Wild-type Lys substituted by Asp"	
FT	Misc-difference 397 /note= "Wild-type Lys substituted by Asp"	
FT	Misc-difference 425 /note= "Wild-type Asp substituted by Lys"	
FT	Misc-difference 426 /note= "Wild-type Asp substituted by Lys"	
XX		
XX	WC200182788-A2.	
XX		
PD	08-NOV-2001.	
XX		
XX	04-MAY-2001; 2001WO-US014372.	
XX		
XX	04-MAY-2000; 2000US-0201800P.	
XX		
PA	(HARD) HARVARD COLLEGE.	
XX		
PI	Collier RJ, Sellman BR;	
XX		
XX	WPI; 2002-017725/02.	
XX		
XX	Protecting humans against anthrax using mutant B groups (anthrax	
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus	
PT	anthracis.	
XX		
PS	Claim 4; Page; 77pp; English.	
XX		
CC	The invention relates to antibacterial agents comprising mutant forms of	
CC	pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants	
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B	
CC	moiety is anthrax protective antigen (PA) and using these mutants or	
CC	compositions of them for protecting against Bacillus anthracis infections	
CC	in humans, especially as vaccines. Note: The present sequence is not	
CC	given in the specification but is derived from the Bacillus anthracis	
CC	wild-type PA protein sequence shown in figure 13 (AAM51483)	
XX		
SQ	Sequence 735 AA;	
	Query Match 99.5%; Score 3753; DB 5; Length 735;	
	Best Local Similarity 99.6%; Pred. No. 1.6e-247;	
	Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	1 EVKQENRLNSESSESSQGLLGYYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60	
Db	1 EVKQENRLNSESSESSQGLLGYYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60	
QY	61 QSAIWSGFTKVKSDYEYFATSADNHVTMWDDQEVINKASNSKIRLEKRLYQIKIY 120	

Db 541 QYQKDIETDFDNFQDOTSONIKNQLAELNATNIYTVLDKIKLNKAKWMLIRDKRFHYDR 600
QY 601 NNIAVGADESUVKEAHRVINSSTGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHRVINSSTGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 11
AAM51483
ID AAM51483 standard; protein; 735 AA.
XX AC AAM51483;
DT 01-FEB-2002 (first entry)
DE Anthrax PA protein.
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.
XX OS Bacillus anthracis.
XX PN WO200182788-A2.
XX PD 08-NOV-2001.
XX PF 04-MAY-2001; 2001WO-US014372.
XX PR 04-MAY-2000; 2000US-0201800P.
XX PA (HARD) HARVARD COLLEGE.
XX PI Collier RJ, Sellman BR;
XX DR WPI; 2002-017725/02.
XX DR N-PSDB; AAI99904.
XX PT Protecting humans against anthrax using mutant B groups (anthrax
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
PT anthracis.
XX PS Disclosure; Fig 13; 77pp; English.
XX CC The invention relates to antibacterial agents comprising mutant forms of
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. The present sequence is that of the
CC anthrax PA protein
XX SQ Sequence 735 AA;

Query Match 99.5%; Score 3753; DB 5; Length 735;
Best Local Similarity 99.6%; Pred. No. 1.6e-247;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNESESSQGLGYFSDLNFPQAPMVVTSSTGDLSTPSELENIPSENOYF 60
Db 1 EVKQENRLNESESSQGLGYFSDLNFPQAPMVVTSSTGDLSTPSELENIPSENOYF 60
QY 61 QSAINSGFIKVKSDBYTTFATSADNHVTWVDDQEVINKASNSNKIRLEKRLYQIKIQY 120
Db 61 QSAINSGFIKVKSDBYTTFATSADNHVTWVDDQEVINKASNSNKIRLEKRLYQIKIQY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSDDNLOQLPELKQSSNSRKRKSTSSAGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSDDNLOQLPELKQSSNSRKRKSTSSAGPTVPDRDN 180
QY 181 DGI-PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWKSTASDPYSDPEKVT 240
Db 181 DGI-PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWKSTASDPYSDPEKVT 240
QY 241 GRIDKNVSPPEARPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTGRHT 300
Db 241 GRIDKNVSPPEARPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTGRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAFSSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAFSSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPPSKNLAPIO 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPPSKNLAPIO 420
QY 421 LNAQKDCASSTPITMNYNQFLELEKTKQLRLDTPQVYGNIAFYNGRVRVDTGSNWSEV 480
Db 421 LNAQKDCASSTPITMNYNQFLELEKTKQLRLDTPQVYGNIAFYNGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENGNL 540
QY 541 QYQKDIETDFDNFQDOTSONIKNQLAELNATNIYTVLDKIKLNKAKWMLIRDKRFHYDR 600
Db 541 QYQKDIETDFDNFQDOTSONIKNQLAELNATNIYTVLDKIKLNKAKWMLIRDKRFHYDR 600
QY 601 NNIAVGADESUVKEAHRVINSSTGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHRVINSSTGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 12
AAM51488
ID AAM51488 standard; protein; 735 AA.
XX AC AAM51488;
DT 01-FEB-2002 (first entry)
DE Anthrax PA mutant D425A.
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX OS Bacillus anthracis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 425 /note="Wild-type Asp substituted by Ala"
XX WO200182788-A2.
XX PD 08-NOV-2001.
XX PF 04-MAY-2001; 2001WO-US014372.
XX PR 04-MAY-2000; 2000US-0201800P.

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKRSTAGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKRSTAGPTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDKKRTFLSPWISNIHKKGLTKYKSSPEKSWTASDPYDFEKVT 240
Db 181 DGIPDSLEVEGYTVDKKRTFLSPWISNIHKKGLTKYKSSPEKSWTASDPYDFEKVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNNVLTPTSLVGNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVLTPTSLVGNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSASTPTITMNYNQFLEKTKQLRLDTPQVYGNATYVNGRVRVDTGNNSEV 480
Db 421 LNAQKDSASTPTITMNYNQFLEKTKQLRLDTPQVYGNATYVNGRVRVDTGNNSEV 480
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIETDFDNFQDQSTQNTKNOELAEALNATYVLDKIKLAKMNLIRDKRPHYDR 600
Db 541 QYQKDIETDFDNFQDQSTQNTKNOELAEALNATYVLDKIKLAKMNLIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
QY 661 DMLNLSLRQDGKTFIDFKKYNDKPLXYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
Db 661 DMLNLSLRQDGKTFIDFKKYNDKPLXYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 10
AAR60179
ID AAR60179 standard; protein; 735 AA.
XX
AC AAR60179;
XX
DT 25-MAR-2003 (revised)
DT 03-APR-1995 (first entry)
XX
XX
DE Protective antigen of Bacillus anthracis.
XX
KW Anthrax; Bacillus anthracis; fusion protein; protective antigen;
KW protective antigen; cell killing; targeting; toxin; pathogen;
KW intracellular; HIV; human immunodeficiency virus; toxin.
OS Bacillus anthracis.
XX
PN W09418332-A2.
XX

PD 19-AUG-1994.
XX
PF 14-FEB-1994; 94WO-US001624.
XX
PR 12-FEB-1993; 93US-00021601.
PR 25-JUN-1993; 93US-00082849.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Leppla SH, Klimpel K, Arora N, Singh Y, Nichols PJ;
XX WPI; 1994-279753/34.
XX N-PSDB; AAQ70180.
PT Nucleic acid encoding anthrax toxin fusion protein - useful for
PT targeting toxin to specific cells, eg for killing tumour cells or HIV-
PT infected cells.
XX
PS Disclosure; Page 81-83; 124pp; English.
XX
CC The sequence encoding the protective antigen of Bacillus anthracis may be
CC used in the construction of a nucleic acid which encodes a fusion protein
CC comprising the anthrax protective antigen binding domain of the native
CC anthrax lethal factor and a sequence encoding an activity inducing domain
CC of a second protein. The fusion proteins are useful for the specific
CC killing of tumour cells or the killing of cells infected with
CC intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 735 AA;
Query Match 99.5%; Score 3753; DB 2; Length 735;
Best Local Similarity 99.6%; Pred. No. 1.6e-247;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKRSTAGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKRSTAGPTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDKKRTFLSPWISNIHKKGLTKYKSSPEKSWTASDPYDFEKVT 240
Db 181 DGIPDSLEVEGYTVDKKRTFLSPWISNIHKKGLTKYKSSPEKSWTASDPYDFEKVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNNVLTPTSLVGNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVLTPTSLVGNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSASTPTITMNYNQFLEKTKQLRLDTPQVYGNATYVNGRVRVDTGNNSEV 480
Db 421 LNAQKDSASTPTITMNYNQFLEKTKQLRLDTPQVYGNATYVNGRVRVDTGNNSEV 480
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIETDFDNFQDQSTQNTKNOELAEALNATYVLDKIKLAKMNLIRDKRPHYDR 600

FF	Key	Location/Qualifiers	
FT	Misc-difference 427	/note= "Wild-type Phe substituted by Lys"	
FF			
XX	PN	WO200182788-A2.	
XX			
PD		08-NOV-2001.	
XX			
FF		04-MAY-2001; 2001WO-US014372.	
XX			
PR		04-MAY-2000; 2000US-0201800P.	
XX			
PA		(HARD) HARVARD COLLEGE.	
XX			
PI		Collier RJ, Sellman BR;	
XX			
DR		WPI; 2002-017725/02.	
XX			
PT		Protecting humans against anthrax using mutant B groups (anthrax	
PT		protective antigens) of the pore-forming binary A-B toxin of Bacillus	
PT		anthracis.	
XX			
PS		Claim 4; Page; 77pp; English.	
XX			
CC		The invention relates to antibacterial agents comprising mutant forms of	
CC		pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants	
CC		in the B moiety of the pore-forming binary A-B anthrax toxin, where the B	
CC		moiety is anthrax protective antigen (PA) and using these mutants or	
CC		compositions of them for protecting against Bacillus anthracis infections	
CC		in humans, especially as vaccines. Note: The present sequence is not	
CC		given in the specification but is derived from the Bacillus anthracis	
CC		wild-type PA protein sequence shown in figure 13 (AAM51483)	
XX			
SQ		Sequence 735 AA;	
	Query Match	99.5%; Score 3754; DB 5; Length 735;	
	Best Local Similarity	99.6%; Pred. No. 1.4e-247;	
	Matches 732; Conservative	0; Mismatches 3; Indels 0; Gaps 0;	
QY	1	EVKQENRLNESSSQGLGYFSDLNFPQPMVVTSSTTGDLSPSSSELENIPSENQYF 60	
DB	1	EVKQENRLNESSSQGLGYFSDLNFPQPMVVTSSTTGDLSPSSSELENIPSENQYF 60	
QY	61	QSAINSGPDKVKSDEYTFATGADNHVTWVDQEVINKASNSKIRLEKGLVQIKQY 120	
DB	61	QSAINSGPDKVKSDEYTFATGADNHVTWVDQEVINKASNSKIRLEKGLVQIKQY 120	
QY	121	QRENPEKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSNRSRKRSTASGPTVDRDN 180	
DB	121	QRENPEKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSNRSRKRSTASGPTVDRDN 180	
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWSSTASDPYDFEKT 240	
DB	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWSSTASDPYDFEKT 240	
QY	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300	
DB	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300	
QY	301	SEVHGNAEYHAFDFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
DB	301	SEVHGNAEYHAFDFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
QY	361	NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420	
DB	361	NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420	
QY	421	LNAQKQASPTPTMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVDVTGNSWSEV 480	
DB	421	LNAQKQASPTPTMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVDVTGNSWSEV 480	
QY	481	LPOIQETTARIIFNGKDLNVERRIAANVPSPDLETTKPDMTLKEALKIAPGFNEPNGNL 540	

481 LPQIQETTARIIFNGKDLNVERRIAANVPSPDPLETTKPDMTLKEALKIAFGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSPDPLETTKPDMTLKEALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFQOQTSQNTKQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFQOQTSQNTKQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDKIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGTTFIDPKYNDKILPLYISNPYKNNYAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGTTFIDPKYNDKILPLYISNPYKNNYAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 7

AAW51487
ID AAW51487 standard; protein; 735 AA.

XX AC AAW51487;

XX 01-FEB-2002 (first entry)

XX Anthrax PA mutant K397Q.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX Bacillus anthracis.

XX Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 397 /note= "Wild-type Lys substituted by Gln"

XX W0200182788-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

XX (HARD) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPT; 2002-017725/02.

XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.

XX Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAW5113 and AAW51484-AAW51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAW51483)

XX Sequence 735 AA;

Query Match 99.5%; Score 3754; DB 5; Length 735;

Best Local Similarity 99.6%; Pred. No. 1.4e-247; Indels 0; Gaps 0;
Matches 732; Conservative 0; Mismatches 3;

QY 1 EVKQENRLNSESSESSQGLLYGFSDLPFQAPMVTSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFSDLPFQAPMVTSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWGSFIVKVKSDSYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWGSFIVKVKSDSYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKEGLDFKLWTDSONKKEVISSDNLQLPKQKSSNSRKRSTTSAGTTPVDRDN 180
Db 121 QRENPTKEGLDFKLWTDSONKKEVISSDNLQLPKQKSSNSRKRSTTSAGTTPVDRDN 180
QY 181 DGIPDSLEVEGYTVDKNRKTFELSPWISNIHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240
Db 181 DGIPDSLEVEGYTVDKNRKTFELSPWISNIHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNNVLTPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVLTPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTITMNNYQFLEKTKQLRLDTPQVYGNATYVNFNGRVRVDTGSNWSEV 480
Db 421 LNAQKDASSPTITMNNYQFLEKTKQLRLDTPQVYGNATYVNFNGRVRVDTGSNWSEV 480
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSPDPLETTKPDMTLKEALKIAFGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSPDPLETTKPDMTLKEALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFQOQTSQNTKQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFQOQTSQNTKQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDKIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGTTFIDPKYNDKILPLYISNPYKNNYAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGTTFIDPKYNDKILPLYISNPYKNNYAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 8

AAW51500
ID AAW51500 standard; protein; 735 AA.

XX AC AAW51500;

XX 01-FEB-2002 (first entry)

XX Anthrax PA F427K mutant.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX Bacillus anthracis.
XX Synthetic.

```
XX SQ Sequence 735 AA;
Query Match 99.7%; Score 3759; DB 5; Length 735;
Best Local Similarity 99.7%; Pred. No. 6 4e-248;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
DB 1 EVKQENRLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDQEVINKASNSNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDQEVINKASNSNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKKRSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKKRSTAGTVPDRDN 180

QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSDPEKVT 240
DB 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSDPEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDASSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYVNFENGRVVDTSNWSSEV 480
DB 421 LNAQKDASSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYVNFENGRVVDTSNWSSEV 480

QY 481 LPQIQETTARIIPNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFQNEPNGNL 540
DB 481 LPQIQETTARIIPNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFQNEPNGNL 540

QY 541 QYQKDIITDFDNFDQOOTSQNIKNQLAELNANIVTVLDKIKLNAMNILLRDKRPFHYDR 600
DB 541 QYQKDIITDFDNFDQOOTSQNIKNQLAELNANIVTVLDKIKLNAMNILLRDKRPFHYDR 600

QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660

QY 661 DMLNISLRQDGKTFIDFKKYNKPLIYISNPYKVNVAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNISLRQDGKTFIDFKKYNKPLIYISNPYKVNVAVTKENTIINPSENGDTSTNG 720

QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 6
AAW51490
ID AAW51490 standard; protein; 735 AA.
AC AAW51490;
XX
DT 01-FEB-2002 (first entry)
DE Anthrax PA mutant D425E.
XX
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX
```

```
OS Bacillus anthracis.
OS Synthetic.
FX Key Location/Qualifiers
FH Misc-difference 425
FT /note= "Wild-type Asp substituted by Glu"
XX
FN W0200182788-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US014372.
XX 04-MAY-2000; 2000US-0201800P.
XX (HARD ) HARVARD COLLEGE.
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
PT anthracis.
XX
XX Claim 4; Page; 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
CC pore-forming toxins (AAW51484 and AAW51484-AAW51500), especially mutants
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAW51483)
XX
XX Sequence 735 AA;
Query Match 99.5%; Score 3755; DB 5; Length 735;
Best Local Similarity 99.6%; Pred. No. 1.2e-247;
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
DB 1 EVKQENRLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDQEVINKASNSNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDQEVINKASNSNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKKRSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKKRSTAGTVPDRDN 180

QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSDPEKVT 240
DB 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSDPEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDASSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYVNFENGRVVDTSNWSSEV 480
DB 421 LNAQKDASSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYVNFENGRVVDTSNWSSEV 480
```

KW	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;	
KW	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.	
XX	Bacillus anthracis.	
OS	Synthetic.	
XX	Key	Location/Qualifiers
XX	Key	Location/Qualifiers
FT	Misc-difference 425	/note= "Wild-type Asp substituted by Lys"
FT	Misc-difference 425	/note= "Wild-type Asp substituted by Lys"
XX	WO200182788-A2.	
XX	08-NOV-2001.	
XX	04-MAY-2001; 2001WO-US014372.	
XX	04-MAY-2000; 2000US-0201800P.	
XX	PA (HARD) HARVARD COLLEGE.	
XX	Collier RJ, Sellman BR;	
XX	WPI; 2002-017725/02.	
XX	Protecting humans against anthrax using mutant B groups (anthrax	
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus	
PT	anthracis.	
XX	Claim 4; Page; 77pp; English.	
XX	The invention relates to antibacterial agents comprising mutant forms of	
CC	pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants	
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B	
CC	moiety is anthrax protective antigen (PA) and using these mutants or	
CC	compositions of them for protecting against Bacillus anthracis infections	
CC	in humans, especially as vaccines. Note: The present sequence is not	
CC	given in the specification but is derived from the Bacillus anthracis	
CC	wild-type PA protein sequence shown in figure 13 (AAM51483)	
XX	Sequence 735 AA;	
SQ	Query Match	99.7%; Score 3759; DB 5; Length 735;
	Best Local Similarity	99.7%; Pred. No. 6.4e-248;
	Matches 733; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1 EVKQENLLNESSSQGLGYVPSDLNFOAPMVVTSSTGDLSTPSSSELENIPSENQYF 60	
DB	1 EVKQENLLNESSSQGLGYVPSDLNFOAPMVVTSSTGDLSTPSSSELENIPSENQYF 60	
QY	61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120	
DB	61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120	
QY	121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNSRKRSTAGTVPDRDN 180	
DB	121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNSRKRSTAGTVPDRDN 180	
QY	181 DGIPDSLEVEGYTVDVKNKTEFLSPWISNIHEKKGLTKYKSPKWKSTADSPYDFEYVT 240	
DB	181 DGIPDSLEVEGYTVDVKNKTEFLSPWISNIHEKKGLTKYKSPKWKSTADSPYDFEYVT 240	
QY	241 GRIDKNVSPARPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300	
DB	241 GRIDKNVSPARPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300	
QY	301 SEVHGNAEVHASFDFIGGSVAGFSNSGSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360	
DB	301 SEVHGNAEVHASFDFIGGSVAGFSNSGSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360	
QY	361 NANIRYVNTGTAPIYVNLPTTSVLGKNQTLATIKADENQLSGOILAPNNYPSKNLAPIA 420	
DB	361 NANIRYVNTGTAPIYVNLPTTSVLGKNQTLATIKADENQLSGOILAPNNYPSKNLAPIA 420	
QY	421 LNAQKASSPTIMYNNQFLEKTKQLRLDQVYGNLATYNFENGRVVRVDTGSNWSEV 480	
DB	421 LNAQKASSPTIMYNNQFLEKTKQLRLDQVYGNLATYNFENGRVVRVDTGSNWSEV 480	
QY	481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGNEPENG 540	
DB	481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGNEPENG 540	
QY	541 QYQKQDITEPDEFNPDQOTSONIKQLAELNATNIVTVLDKIKLNAKWNILIRDKRPHYDR 600	
DB	541 QYQKQDITEPDEFNPDQOTSONIKQLAELNATNIVTVLDKIKLNAKWNILIRDKRPHYDR 600	
QY	601 NNIAVGADSVVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660	
DB	601 NNIAVGADSVVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660	
QY	661 DMLNSSLRODGKTFIDFKKNDKPLYSNPNKYNVAVTKENTIIINPSENGDSTNG 720	
DB	661 DMLNSSLRODGKTFIDFKKNDKPLYSNPNKYNVAVTKENTIIINPSENGDSTNG 720	
QY	721 IKKILIFSCKGYEIG 735	
DB	721 IKKILIFSCKGYEIG 735	
RESULT 5		
AAM51492		
ID	AAM51492 standard; protein; 735 AA.	
AC	AAM51492;	
XX	01-FEB-2002 (first entry)	
XX	Anthrax PA mutant F427A.	
DE	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;	
XX	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.	
XX	Bacillus anthracis.	
OS	Synthetic.	
XX	Key	Location/Qualifiers
FT	Misc-difference 427	/note= "Wild-type Phe substituted by Ala"
FT	Misc-difference 427	/note= "Wild-type Phe substituted by Ala"
XX	WO200182788-A2.	
XX	08-NOV-2001.	
XX	04-MAY-2001; 2001WO-US014372.	
XX	04-MAY-2000; 2000US-0201800P.	
XX	(HARD) HARVARD COLLEGE.	
XX	Collier RJ, Sellman BR;	
XX	WPI; 2002-017725/02.	
XX	Protecting humans against anthrax using mutant B groups (anthrax	
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus	
PT	anthracis.	
XX	Claim 4; Page; 77pp; English.	
XX	The invention relates to antibacterial agents comprising mutant forms of	
CC	pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants	
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B	
CC	moiety is anthrax protective antigen (PA) and using these mutants or	
CC	compositions of them for protecting against Bacillus anthracis infections	
CC	in humans, especially as vaccines. Note: The present sequence is not	
CC	given in the specification but is derived from the Bacillus anthracis	
CC	wild-type PA protein sequence shown in figure 13 (AAM51483)	
XX	Sequence 735 AA;	

Db 361 NANIRYVNTGTAPIYVNLFTTSLVGLKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420
 QY 421 LNAQKDSSTPTIMVYNQFLEKTKQLRLDQVYGNATYVFNENGRVVRVDTGSNWSEV 480
 Db 421 LNAQKDSSTPTIMVYNQFLEKTKQLRLDQVYGNATYVFNENGRVVRVDTGSNWSEV 480
 QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
 Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
 QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAMNILLIRKPHYDR 600
 Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAMNILLIRKPHYDR 600
 QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEEDTEGLKEVINDRY 660
 Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEEDTEGLKEVINDRY 660
 QY 661 DMLNSSLRODQGTFFDFKYNKDLPLYISNPNYKVVAVTKENTIINPSENGDTSTNG 720
 Db 661 DMLNSSLRODQGTFFDFKYNKDLPLYISNPNYKVVAVTKENTIINPSENGDTSTNG 720
 QY 721 IKKILIFSXXGYEIG 735
 Db 721 IKKILIFSXXGYEIG 735

RESULT 3

AAM51485
 ID AAM51485 standard; protein; 735 AA.
 AC AAM51485;

DT 01-FEB-2002 (first entry)

DE Anthrax PA mutant K397D.

KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

KM B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; muten.

OS Bacillus anthracis.

OS Synthetic.

Key Location/Qualifiers

FT Misc-difference 397

FT /note= "Wild-type Lys substituted by Asp"

XX WC200182788-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

XX (HARD) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

XX Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.

XX Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections

CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)
 XX
 SQ Sequence 735 AA;

Query Match 99.7%; Score 3760; DB 5; Length 735;
 Best Local Similarity 99.7%; Pred. No. 5.4e-248;
 Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
 Db 1 EVKQENLLNESSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
 QY 61 QSAIWSGFIVKKSDEYTPATSNADNHVVMVDQEVINKASNSKIRLEKRYIQIKIY 120
 Db 61 QSAIWSGFIVKKSDEYTPATSNADNHVVMVDQEVINKASNSKIRLEKRYIQIKIY 120
 QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQIPKQKSSNRKKESTAGPTVPORDN 180
 Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQIPKQKSSNRKKESTAGPTVPORDN 180
 QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
 Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
 QY 241 GRIDKNVSPARHPLVAAYPVHVDMENIILSKNEOSTONTSETPTISKNTSTSTHT 300
 Db 241 GRIDKNVSPARHPLVAAYPVHVDMENIILSKNEOSTONTSETPTISKNTSTSTHT 300
 QY 301 SEVEGNAEVEHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
 Db 301 SEVEGNAEVEHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
 QY 361 NANIRYVNTGTAPIYVNLFTTSLVGLKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420
 Db 361 NANIRYVNTGTAPIYVNLFTTSLVGLKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420
 QY 421 LNAQKDSSTPTIMVYNQFLEKTKQLRLDQVYGNATYVFNENGRVVRVDTGSNWSEV 480
 Db 421 LNAQKDSSTPTIMVYNQFLEKTKQLRLDQVYGNATYVFNENGRVVRVDTGSNWSEV 480
 QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
 Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
 QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAMNILLIRKPHYDR 600
 Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAMNILLIRKPHYDR 600
 QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEEDTEGLKEVINDRY 660
 Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEEDTEGLKEVINDRY 660
 QY 661 DMLNSSLRODQGTFFDFKYNKDLPLYISNPNYKVVAVTKENTIINPSENGDTSTNG 720
 Db 661 DMLNSSLRODQGTFFDFKYNKDLPLYISNPNYKVVAVTKENTIINPSENGDTSTNG 720
 QY 721 IKKILIFSXXGYEIG 735
 Db 721 IKKILIFSXXGYEIG 735

RESULT 4

AAM51491
 ID AAM51491 standard; protein; 735 AA.

XX AC AAM51491;

XX DT 01-FEB-2002 (first entry)

XX DE Anthrax PA mutant D425K.

XX

CC	compositions of them for protecting against Bacillus anthracis infections
CC	in humans, especially as vaccines. Note: The present sequence is not
CC	given in the specification but is derived from the Bacillus anthracis
CC	wild-type PA protein sequence shown in figure 13 (AAM51483)
XX	
XX	Sequence 735 AA;
QY	Query Match 100.0%; Score 3772; DB 5; Length 735;
DB	Best Local Similarity 100.0%; Pred. No. 8.2e-249;
DB	Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 EVKQENRLNSESSESSQGLGYYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENOYF 60
DB	1 EVKQENRLNSESSESSQGLGYYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENOYF 60
QY	61 QSAIWSGFIKVKSDSEYTFATSDNHNVTMWVDDQEVINKASNKKIRLEKGLYQIKIY 120
DB	61 QSAIWSGFIKVKSDSEYTFATSDNHNVTMWVDDQEVINKASNKKIRLEKGLYQIKIY 120
QY	121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKKGSTAGTPVDPDRN 180
DB	121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKKGSTAGTPVDPDRN 180
QY	181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSTASDPYDFEYVT 240
DB	181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSTASDPYDFEYVT 240
QY	241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB	241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY	301 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
DB	301 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
QY	421 LNAQKDASSPTITMNYNQFLEKTKQLRDTDOVYGNATYNFENGVRVDTGNSWSEV 480
DB	421 LNAQKDASSPTITMNYNQFLEKTKQLRDTDOVYGNATYNFENGVRVDTGNSWSEV 480
QY	481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540
DB	481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540
QY	541 QYQKDIETDFNPDQOTSQNIKNQLAELNATNIYTVLDKIKLNAKNTILIRDKRPHYDR 600
DB	541 QYQKDIETDFNPDQOTSQNIKNQLAELNATNIYTVLDKIKLNAKNTILIRDKRPHYDR 600
QY	601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB	601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY	661 DMLNISRQDGKTFIDPKYNDKPLIYISNPYKVNVAVTKEIINPSENGDSTNG 720
DB	661 DMLNISRQDGKTFIDPKYNDKPLIYISNPYKVNVAVTKEIINPSENGDSTNG 720
QY	721 IKKILIFSKGYEIG 735
DB	721 IKKILIFSKGYEIG 735

RESULT 2
AAM51493
ID AAM51493 standard; protein; 735 AA.

XX AAM51493;

XX 01-FEB-2002 (first entry)

DE Anthrax PA mutant K397D/D425K.

XX	Key Location/Qualifiers
KW	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; murein.
OS	Bacillus anthracis.
OS	Synthetic.
XX	
FT	Misc-difference 397 /note= "Wild-type Lys substituted by Asp"
FT	Misc-difference 425 /note= "Wild-type Asp substituted by Lys"
FT	
PN	WO200182788-A2.
XX	
PD	08-NOV-2001.
XX	
PF	04-MAY-2001; 2001WO-US014372.
XX	
PR	04-MAY-2000; 2000US-0201800P.
XX	
PA	(HARD) HARVARD COLLEGE.
XX	
PI	Collier RJ, Sellman BR;
XX	
DR	WPI; 2002-017725/02.
XX	
PT	Protecting humans against anthrax using mutant B groups (anthrax
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus
PT	anthracis.
XX	
PS	Claim 4; Page; 77pp; English.
XX	
CC	The invention relates to antibacterial agents comprising mutant forms of
CC	pore-forming toxins (AAM5113 and AAM51484-AAM51500), especially mutants
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC	moiety is anthrax protective antigen (PA) and using these mutants or
CC	compositions of them for protecting against Bacillus anthracis infections
CC	in humans, especially as vaccines. Note: The present sequence is not
CC	given in the specification but is derived from the Bacillus anthracis
CC	wild-type PA protein sequence shown in figure 13 (AAM51483)
XX	
QY	Sequence 735 AA;
DB	Query Match 99.8%; Score 3766; DB 5; Length 735;
DB	Best Local Similarity 99.8%; Pred. No. 2.1e-248;
DB	Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 EVKQENRLNSESSESSQGLGYYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENOYF 60
DB	1 EVKQENRLNSESSESSQGLGYYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENOYF 60
QY	61 QSAIWSGFIKVKSDSEYTFATSDNHNVTMWVDDQEVINKASNKKIRLEKGLYQIKIY 120
DB	61 QSAIWSGFIKVKSDSEYTFATSDNHNVTMWVDDQEVINKASNKKIRLEKGLYQIKIY 120
QY	121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKKGSTAGTPVDPDRN 180
DB	121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKKGSTAGTPVDPDRN 180
QY	181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSTASDPYDFEYVT 240
DB	181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSTASDPYDFEYVT 240
QY	241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB	241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY	301 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
DB	301 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
QY	361 NANIRYVNTGTAPIYVLPVTTSLVGNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB	361 NANIRYVNTGTAPIYVLPVTTSLVGNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY	421 LNAQKDASSPTITMNYNQFLEKTKQLRDTDOVYGNATYNFENGVRVDTGNSWSEV 480
DB	421 LNAQKDASSPTITMNYNQFLEKTKQLRDTDOVYGNATYNFENGVRVDTGNSWSEV 480
QY	481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540
DB	481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540
QY	541 QYQKDIETDFNPDQOTSQNIKNQLAELNATNIYTVLDKIKLNAKNTILIRDKRPHYDR 600
DB	541 QYQKDIETDFNPDQOTSQNIKNQLAELNATNIYTVLDKIKLNAKNTILIRDKRPHYDR 600
QY	601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB	601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY	661 DMLNISRQDGKTFIDPKYNDKPLIYISNPYKVNVAVTKEIINPSENGDSTNG 720
DB	661 DMLNISRQDGKTFIDPKYNDKPLIYISNPYKVNVAVTKEIINPSENGDSTNG 720
QY	721 IKKILIFSKGYEIG 735
DB	721 IKKILIFSKGYEIG 735

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:26:27 ; Search time 49.1208 Seconds
(without alignments)
4227.791 Million cell updates/sec

Title: US-09-848-909a-13

Perfect score: 3772

Sequence: 1 EVKQNRLLNSESSESSQGLL.....TSTNGIKLIFSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3772	100.0	735	5	AAM51495 Anthrax P
2	3766	99.8	735	5	AAM51493 Anthrax P
3	3760	99.7	735	5	AAM51485 Anthrax P
4	3759	99.7	735	5	AAM51491 Anthrax P
5	3759	99.7	735	5	AAM51492 Anthrax P
6	3755	99.5	735	5	AAM51490 Anthrax P
7	3754	99.5	735	5	AAM51487 Anthrax P
8	3754	99.5	735	5	AAM51500 Anthrax P
9	3754	99.5	735	5	AAM51489 Anthrax P
10	3753	99.5	735	2	AAR60179 Protective
11	3753	99.5	735	5	AAM51483 Anthrax P
12	3753	99.5	735	5	AAM51488 Anthrax P
13	3753	99.5	735	5	AAM51494 Anthrax P
14	3753	99.5	735	5	AAM51499 Anthrax P
15	3753	99.5	736	3	AAY56959 B. anthra
16	3753	99.5	763	3	AAY56960 B. anthra
17	3753	99.5	764	3	AAY56958 B. anthra
18	3753	99.5	764	4	AAB47306 Wild type
19	3753	99.5	857	7	ADE65872 Bacillus
20	3752	99.5	735	5	AAM51484 Anthrax P
21	3751	99.4	735	5	AAM51486 Anthrax P
22	3750	99.4	735	5	AAM51488 Anthrax P
23	3750	99.4	764	6	AAB35717 Bacillus
24	3738	99.1	764	6	ABP71693 B. anthra
25	3708.5	98.3	903	2	AAR60183 PA(1-725)

ALIGNMENTS

RESULT 1

AAM51495

ID AAM51495 standard; protein; 735 AA.

XX AC AAM51495;

XX DT 01-FEB-2002 (first entry)

XX DE Anthrax PA mutant K397D/D425K/F427A.

XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX OS Bacillus anthracis.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"

FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"

FT Misc-difference 427 /note= "Wild-type Phe substituted by Ala"

XX WO200182786-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

XX (HARD) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

XX Protecting humans against anthrax using mutant B groups (anthrax

XX protective antigens) of the pore-forming binary A-B toxin of Bacillus

XX anthracis.

XX Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of

XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants

XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B

XX moiety is anthrax protective antigen (PA) and using these mutants or

AAM51498 Anthrax P
AAM51497 Anthrax P
AAM51496 Anthrax P
AAM52113 Anthrax P
AAR60193 Modified
AAE18288 Bacillus
AAY56961 B. anthra
AAE18285 Bacillus
AAE18287 Bacillus
AAE18286 Bacillus
AAE18284 Bacillus
AAM50707 Bacillus
AAE07903 C. botuli
AAE07901 C. botuli
AAE07900 C. botuli
AAE07902 C. botuli
AAE35719 Clostridi
AAW60224 Bacillus
AAY59277 MIS toxin
AAR91239 B. cereus

26 3626 96.1 711 5 AAM51498
27 3620 96.0 711 5 AAM51497
28 3613 95.8 711 5 AAM51496
29 3607 95.6 711 5 AAM52113
30 3576 94.8 719 2 AAR60193
31 3037 90.5 595 5 AAE18288
32 2898 76.8 569 3 AAY56961
33 2488 66.0 487 5 AAE18285
34 2173 57.6 426 5 AAE18287
35 1624 43.1 318 5 AAE18286
36 1338 35.5 258 5 AAE18284
37 1275 33.8 288 5 AAM50707
38 839.5 22.3 1052 4 AAE07903
39 838 22.2 1032 4 AAE07901
40 834 22.1 1092 4 AAE07900
41 825.5 21.9 1112 4 AAE07902
42 810 21.5 721 6 AAE35719
43 767.5 20.3 880 2 AAW60224
44 767 20.3 881 3 AAY59277
45 766 20.3 884 2 AAR91239

Db 1395 NKVHNFINSYFDLNEKKNIF 1419

Search completed: May 3, 2004, 19:40:03
Job time : 38.6261 secs

```
Db 942 LLSITSYSSYGGKKKELMIKYEKTLKANTRIQLK-NWKNFQKQFEMFSIENMNI-----996
Qy 158 QKSSNKKRSTKSAGTVPDRNDGIPDS-LEVEGYTVGVK-RTFLS-----PKI 207
Db 997 -----DIPLDPIFDPIILKIKVKKYKSGYFGETNISLVYLPWI 1037
Qy 208 SNIHEKKGLTKYKSSPEKSTASDPYSDPEKV-----TGRIDKNSPEARHPLVAAYPIV 262
Db 1038 KNIDE-----VLQAHDDYSEIINMKNIDTNIYKN-----KNAALVISALS 1082
Qy 263 HVDMENILSKNDQSTQNTDST-----RTISKNTSTSTHSEVHGNEVHASFPD 315
Db 1083 LADCEPTLSKERNKYNKENDDEAWKEIPLFLNDQENQKEDKNKTSQSGHNTYNYDGN 1142
Qy 316 IGGVSAGFNSNSTVAIDHLSLAGERTWAGTGLTADTARLANIRYVNTGTAPT- 374
Db 1143 -NGAYEMGYMETIYAKNDN-----NNNNNNYNNYNNYNNYNNYNYAAPT 1192
Qy 375 -YVLPPTSLVLGKNOTLATIDADENQLSQILAPNYPFSKNLAPIALNAQKFSSTPI 433
Db 1193 SYN-----NNVLQDTRNNVRYNHSNNMNMNNYKNIYNASQFVINYNNYNYDKGNT 1248
Qy 434 MNYN-----QFLEKTK-----QLRLDQDVYGNIAATYN-FENGVRVDTGSNWSEVL 481
Db 1249 LNFNNNNIHFKNLSNNKFSYLSRQKDTYNIKNYNSIYKLPDDGIPILIKLSYVKNY 1308
Qy 482 POIQTETARIIFN-----GKDLNVERRIAAYN-----PSD 512
Db 1309 FYIKILTSKYLNVHIPPRIYLVGDKLN-IEKFIKNIRNVSDGILENLYDDILIPSL 1367
Qy 513 PLETTKEDMTL-----KEALKIA-----FGNEP-NG-----NLQYGGKIDTFD 551
Db 1368 PLIKCNDISCDNNYENKIEKGIFGCFEQFPFVEIIGQIKCFKIKYRNLESENMP 1427
Qy 552 FNEPQOTSONI-----KNOLA-----ELNATNIYTVLDKIKLNAMNILRDK-- 594
Db 1428 LSLKDIITQNIIFRNKFRGKNKIPLYKIRVYVLRGIGLYGINNEYTANPYLIFSLEKTS 1487
Qy 595 --RHYDRNNI-----AVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIETE 647
Db 1488 NLRNAPKRSNINPEFGCLWSEAFPE--DEIITISVYSAEDNYDKOINDIYG-----1539
Qy 648 DTEGLEKEVIN--DRYDMLNISSLRQDGKTFIDFKK-YN--KLPLYISNPNYKV-----N 697
Db 1540 STE-----INLFDWMSKEWRHMKKNKIPVEYRPLYSNYIKHPKVSSNNYNTMWSNN 1594
Qy 698 VYAV-----TKENTINPSENGDTS 717
Db 1595 IFSPFDIFNYLMTYTSPTKGNNNNNNDNNNS 1627

RESULT 15
Q8IHP9 PRELIMINARY; PRT; 2940 AA.
AC Q8IHP9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF11_0480.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
EX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
```

```
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Murgall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrrell B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014843; AAN36060.1; .
KW Hypothetical protein.
SQ SEQUENCE 2940 AA; 348040 MW; BF8748D8A051BAD2 CRC64;
```

```
Query Match 5.28; Score 196.5; DB 5; Length 2940;
Best Local Similarity 18.78; Pred. No. 1.2; Indels 243; Gaps 40;
Matches 162; Conservative 137; Mismatch 323;
```

```
Qy 34 VVTSSTTGDLSIPSELENIPSENQYFQSAIWSGFIKVKKSDYTFATSDADNHYTMVDD 93
Db 627 IINHNNDISINHNHN-----NYIFMNYMNNI-----NNYCYCYMNNNTHVANNYN 679
Qy 94 QEVINKASNK-----IRLEKGLYQIKIYOQRENETEGLDFK-- 133
Db 680 IYIQNHQDNAPILQINNHIAHINDLCYHSEKNEYTKISKNHQMNINPOQSGNQ 739
Qy 134 -----LYWTDSONK-----KEVISSNQLQPELKQKSSNRKRSAGPT-- 174
Db 740 NDISNNINKDEYVNLNEQKILCDKNKSIKCDIPQKCDNTQDENSEQONVITPNSG 799
Qy 175 -----VPDRNDGIP--DSLEVEGYTVGVKKTFLSPWISNHEKKGL----- 216
Db 800 HYKIQEQNNYVQHIFDEHNTTNEIMTQNT-----NINLDPYIMNQNNDNNVYLNQ 853
Qy 217 ----TKYKSPE--KWTASDPYSD----FEKVTGRIDK--NVSPEARHPLVAAYPIVHVD 265
Db 854 NYFTEKKEEDINITEANTHYNNVHIYQNNSHHNMKNLNNNTSKSTQDYDLVHNN 913
Qy 266 ME-----NIIISK--NEDQST-----QNTDSETTISKVTS--T 295
Db 914 MENFNNTYINIMKEGIHNDTSSVINSSINQNLCKNKKTNDYQNRNINKNDSIPD 973
Qy 296 SRTTSEVHGNAEVAHGFEDIGGSVAGFNSNSTVAIDHLSLAGERTWAGTGLNTA 355
Db 974 SSINWELKNIQNDY--ASNIYNNNNNNNNIISNNI--ISYNNY-----NNE 1024
Qy 356 DT-ARLAN--IRYVNTGTAPIYVLP--TTSVLGKNOTLATIDADENQLS-----Q 403
Db 1025 NVYFQINSNNYIPVID-----PHLNNPEYKHTINNQLNKNKNCINNISMDDNDYDSNTISIH 1080
Qy 404 ILAPNNYPSKNLAPIALNAQKFSSTPIWVYVYQFLEKTKQLRLDQDVYGNIAATYN 463
Db 1081 ISVNDNEYNNNFNDNMNNNN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1132
Qy 464 FENGVRVDTGSNWSEVLPOIQETTA-RIIFNG-KDLNVERRIAAYNPSDPLETTKPD 521
Db 1133 LYSNNYNIQHPSSNDVPKIQKHIANNIIMNGKKEEHIIEKK-----SKEETNKTE 1185
Qy 522 TLKEALKIAGFNPNGNLQYGGKIDTFEDFNFQOTSONIKNQLAEL-----NA 571
Db 1186 QYRSI-----NQNTLILKENEIDENDINTLQ--NLNFKNDMMNNDNIINLKNLNI 1237
Qy 572 TNYITVLDKIKLNAMNILRDKGFHYDRNNIAVGADESUVKEAHEVINSSTEGLLNI 631
Db 1238 NNIYT-----PYQNNILKNNIQLFNNN-----KEVITKAYTHSSNEININVKNI 1284
Qy 632 KDKIRKILSGYIETEDETEGKIDRYDMNLNLSLRQDGKTF-----ID- 677
Db 1285 DTQI-----NIRKNENNNDIINK-EQKJSTNNEQNTYNTTSSVKNYSILDN 1334
Qy 678 ---PKYNDKLPXYISN-----NYKVNVAVTKEK-----IINPSENG 714
Db 1335 SPTDKYKKNLNLNLSSTSYNTNDINNVHDDIDKTNKKNYEEBKHKDINIYINNNSG 1394
Qy 715 DTSTNGI-----KKILIF 727
```

QY	682	NDKLP	PLV	ISNP	KYV	YVAT	KENT	ILN	SENG	DTS	717
Db	4416	SNKLP	KLANN	YQIK	---	INKS	HTII	---	SKNG	EW	4445
RESULT 13											
Q8IHV4											
ID	Q8IHV4	PRELIMINARY;									
AC	Q8IHV4										
DT	01-MAR-2003	(T-EMBLrel. 23, Created)									
DT	01-MAR-2003	(T-EMBLrel. 23, Last sequence update)									
DT	01-MAR-2003	(T-EMBLrel. 23, Last annotation update)									
DE		Hypothetical protein.									
GN	Pf11_0392.										
OS	Plasmodium falciparum (isolate 3D7).										
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.										
NCBI	Taxid:36329;										
OR	[1]										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=3D7;										
RC	MEDLINE=2255705; PubMed=12368864;										
RA	Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,										
RA	Carlton J.M., Pain A., Nelson K.E., Bowman S., Faulsen I.T., James K.,										
RA	Eisen J.A., Rutherford K., Salzberg S.B., Craig A., Kyes S.,										
RA	Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,										
RA	Pertea M., Allen J., Selengut J.J., Haft D., Mather M.W., Vaidya A.B.,										
RA	Martin D.M.A., Fairlamb A.H., Subramanian M.J., Roos D.S., Ralph S.A.,										
RA	McFadden G.I., Cummings L.H., Fraumonian G.M., Mungall C.,										
RA	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,										
RA	Fraser C.M., Barrell B.,										
RT	"Genome sequence of the human malaria parasite Plasmodium										
RT	falciparum";										
RL	Nature 419:498-511(2002).										
DR	EMBL; AE014841; AAN35975.1; -.										
KW	Hypothetical protein.										
DS											
SQ	SEQUENCE 2849 AA; 335916 MW; B5515D173D96B813 CRC64;										
Query Match											
Best Local Similarity	5.3%; Score 201.5; DB 5; Length 2849;										
Matches	170; Conservative 134; Mismatches 333; Indels 189; Gaps 41;										
QY	3	KQENLLNESSSSQGLGV-YFSDLNFO--APWVTSSTTGPLSP--SSELENTPSN	57								
Db	1862	KHNYTFKKNLKIPIKYNNNNNINEKDDIPKSIQSSFINEDNAYINKEGNNKLN	1921								
QY	58	Q-YFQSAIWGGFIKVKKSDEYFATGADNHVTVMVDD-----QEVINKASNK-	105								
Db	1922	EDIIQERFSHNHKTVMNKNCSYSDCNIVKINYDELNDSTQTKELNKGKSNNGKAEW	1981								
QY	106	-ILEKGRLY-QIKIQYQRENPEKGLDFKLYTWDSONKEVSISSDNLQELPKQKSSNS	163								
Db	1982	IIDIKKNETFPYIKIDKKEKNDEK--KHKYMYKNDKNIMKGSN---KSGMKONKNS	2036								
QY	164	RKRSR-SAGPTVPDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIEKKGLTKYKSSP	223								
Db	2037	NKMKH-----IPLSVNNGY-----NKSSINKKYENNIN-----KNK	2069								
QY	224	EKWSIASDPYSDF--EKVTGRI-DKNVSPFARPLVAAPYIVHDM-----NILS	272								
Db	2070	DKLNILVNSISKLVQSKIKOELSNKNIKD-----ILNFEITIKKKSKKETKNTWHTN	2124								
QY	273	KNEIDQSTQNTDSETRITSKNTSRTHTSEV-HGNAEVHASPFIDIGSVSAG-FSNSNGS	330								
Db	2125	KXNDNNNDNNINFERINDNINSYVNLNKVPYDNI-----DEGNYINGDIYAPDS	2179								
QY	331	TVAIDHLSIAGERTWAETWGLNTA-----DTARLNA-----NIRYVNTGTAPIYNVL	378								
Db	2180	TI-----YTWVNNIDTMYNMYKYPDSKKNINHVPL									

DR EMBL; AE014841; AAN35955.1; -
KW Hypothetical protein.
SQ SEQUENCE 3468 AA; 411913 MW; 506F7D62999BA7B1 CRC64;

Query Match
Best Local Similarity 18.9%; Pred. No. 0.36;
Matches 158; Conservative 130; Mismatches 276; Indels 271; Gaps 37;

QY 55 SENCYFQSAWSGFIKVKKDEYTFATSDADNHYTMVMDQEV---INKASNNKIRLEK 111
DB 160 SENEY-----TKNDCKYNNSTNGY-----NKEIFLKNIAKEHSLKKIEKN 202
QY 112 RLVOIK-----IQVQRENPEKGLDFK---LYM---TDSON-----KKEVTSDDL 151
DB 203 INLLKCDSDNNDLBPKQNVHLNNDVASEDLFHDNRVNNQNLKDIKDYIHHDDNI 262
QY 152 QLPKCKSSNRKRSKTSAGTPVDRDNDGIPDSLEVEGYTVDV-KNKRTTFLSPWISNI 210
DB 263 MI-----ESNDNRNDKCNKFNTPYNERHIIIVDKTKKGNNNINNSNNFDS--ISNI 315
QY 211 HEKGLTKYKSSPEKMWSTADP---YSDFE-KVTGRIDKQVSPPEARHPLVAAYPIVHV- 264
DB 316 NKK-----ISYPINMYNSSEDKTLNQILNDSILSDSLKQAYSPYSL 360
QY 265 ---DMENIILSKEDQSTQTD--SETRTSKNTSTRTSTSEVHGNAEVAH--SPFDIG 317
DB 361 EKNNNGNNSHNNNEVYRDVIDEDINISKFSINLSNLDSEMDNNEFNSLCSP---- 416
QY 318 GSVSAGFSNSSTVAIDHSLSLAGERTWAEETMGL-----NTADTARLANIRY- 366
DB 417 ---NSSKCVKDEMITQFVGNRYKVTMDVANDNYMKNSIEQHSINMESNHFKNQ 468
QY 367 VNTGTAPIYVNLPTTSLVLGKQ-----TLATIDADE--NQLSQILAPNN--- 409
DB 469 KINSKGEDANNLNSHILNNKIGIQVNSIISDYLYSIKKQESWNNNSNNEALNNIV 528
QY 410 -----YYPKSLAPI-----ALNAQKFSSTP-- 431
DB 529 TNNNGSNNKNSNVYKTSQYVYVNPNDNNNNHQLNLSYMKNSNNLNTSNGFKKIPKN 588
QY 432 ---ITMNVNQLFLEKTKQLRLDTPQVYGNIAVTYFENGVRVDTGNSWSEVLPOQ 485
DB 589 KNIISNIDFDNNIFKSYIKENVVKQNSNINHOVEKQNTYNDSEINKN-----NNIENNTQ 644
QY 486 ETTAR-----IIFNGKDLNLVERIAAVNPSDPLETTK----- 518
DB 645 NTTCNFTNTDNDIINKKIKIYKIKIDSISLNDLSLNKLSLNSISMDRYTKNYYEK 704
QY 519 --PDMTLKEAL-----KIAFGNEPQNLQYCGKDITEFDNFQOQTS 559
DB 705 FLDDVILDDSFATSNELLQHSNYTTTWHIFDNNNNNNNNINQKEDLFQNDYN---KET 761
QY 560 QNISKQALAEINATYVLDKIKLNKQNLIRDKRPHYDRNNIAGVADSVVKEAREV 619
DB 762 YNNIMLSNNAENLF-----KISYSCDLVLGKNEIILD--RVV 800
QY 620 INSSTEGLLLNIDKIRKILSGYIETEDTEGLKE-----VINDRYDMLNISS--L 668
DB 801 ENSKTE-----QVEYQNKEDIKLYHKDDNIIDNNYNDVNIKKDCHL 844
QY 669 RDGKTETDFKYN--KLPL-----YISNPNKYV--VYVTKENTILNPSN 713
DB 845 KMDNQDNKINKSQSKKKNPINNNECNVHNKFNKINEIDHPKEENTLNYESKN 899

RESULT 12

Q9PQ08 PRELIMINARY; PRT; 4688 AA.
AC Q9PQ08;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein UU482.

GN UU482.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serovar 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
urealyticum";
RL Nature 407:757-762(2000).
DR EMBL; AE002145; AAF30894.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 4688 AA; 534880 MW; B53ABFAFEE1997E CRC64;

Query Match
Best Local Similarity 5.4%; Score 202; DB 16; Length 4688;
Matches 174; Conservative 131; Mismatches 325; Indels 186; Gaps 40;

QY 19 LLGYFFDLNFPQAPMVYTSSTGDLSPSSSELENIPSENQVFOSAIWSGFIKVKKSEY 78
DB 3699 LVDVYYLD-NIHQIDETRKIFKDHV-SKEIETNPQVWISKHGNNKSPDTTANFEFK 3756
QY 79 FATSADNHYTMVMDQVINKASNNKIRLEKGLYQIKCYORENFTKGLDFKLYWTD 138
DB 3757 IETQ-----DNDVNLNIDATVFKOEHNNIKOKIVRIKEN-----ND 3795
QY 139 SONKEVISSNLQLPKQKSSN-----SRKRSSTAGTPVDRD----- 179
DB 3796 WLIQGI---DNLN-PETKYKLENIELSKPLKTHLSVINDKENISLITETGNPVLKV 3851
QY 180 ---NDGTPDSLEVEGYTVDKNKETFLSPWISNIHEKGLTKYKSSPEKMWSTADSYSD 235
DB 3852 IOTQNTDITNDQOQINVTLSGVNSK-YNGRQIKVYKNNNNVIYESS---LITLQKQND 3907
QY 236 FEKVTGRIDKQVSPPEARHPLVAAYPIVHVDMEINILSKN-EDQSTQNTDSET-RTISKNT 293
DB 3908 YQLLLSNLNSN-----REYRFEKIEINHISNTNPEDLEKLGVSNTFITQKNT 3957
QY 294 STSRTHTS-EVHGNAEVHASP-----FDIGSVSAGFS-----NSNS--STVAID 335
DB 3958 TVQWNSDSATTIVGTVGNFNFKIKSEDKILENNOQVAVFAPKETIRDTNLTWQYRPLK 4017
QY 336 HSLSLAGERTWAEETMGLNT---ADTARLANIRVYVNTGTAPIYVNLPTTSLVLGKNQTLA 392
DB 4018 DVTSDFKGTWAHDLNSVNFKEETTYKLVKIQVFNKETKAKNNINNSNNVILDNTNSI 4077
QY 393 -----TIDADENQLSQILAPNYYPSKQLAPIALNAQKK-FSSTPTTMNY--NOFLELE 443
DB 4078 NSNYEFTTKVGDHKLINITSSNNVTNSQTINFTLSGVKKSWSVGKIKLSYKSDTSESI 4137
QY 444 KTKQLRLDTPQVYGNIAVTYFENGVRVDTGNSWSEVLPOIQTETTARIIENGKD 497
DB 4138 HTNEVLIESNKTQYNILLNLLNKKRNTYTLIDVLIDNNVSDFPKEGNLTNSFITRTSA 4197
QY 498 LNLVERRIAAVNPSDPLETT-----KPDWTLKEA-----LKIAFGNEPQNLQ 541
DB 4198 INVLNIEISNRASTNLKSTIIKINLNDPDNVLKDQQAATTIVYGNKQAMGFIIVSGNIK 4257
QY 542 YQKQDITEFDNFQOQTSQNIK-NOLAEINATYVLDKIKLNKQNLIRDKRPHYD- 599
DB 4258 YLTATVLDLNFN-DKVNIVNISFNKPSIAAEN-----IGDKSNII-----YND 4304
QY 600 -----RNNAVGA---DESIVKEAREVINSSTEGLLNIDKIRKILSGYIETEDTE 650
DB 4305 IPKLEINNDIIVNGPINKKEIVVKVANOK--NNIDVDLGLQINPKIAHLNR-FLAKFKSTN 4361
QY 651 GLKEVIND--RYDMLNITSSL-RDQKTFIDF-----KKY----- 681
DB 4362 -----NDIIEITVNGSSLVNNDGKTSIRFTLNKANKLYSLVDVYLVNNSNTIVE 4415

Db 1 MESLGINNINYNALDRILAKNKNILVRDP-YHYDNGNIVGVDDSYLKNAYQILNWSDD 59
Qy 626 GLLNTDKDIRKILSYVIEIDTE-----GLKEVINDRYDMLNISSLRQDG 672
Db 60 GVSALNDEVDQALSGYMLQIKKPSNHLTNSPVTITLAKDSGVGELYRVLIS-----DG 113
Qy 673 KTFIDPKYNDKLPILVSPNKNVYVYTKEN-TIINPSENGDSTNGICKILIFSKKG 731
Db 114 TQFLDNKFDENKRSVLV-DPGDDVYVAVTKEDFNAVTRDENGNTA-NKIKNTLVLSGKI 171
Qy 732 YEI 734
Db 172 KEI 174
RESULT 10
Q8RGK2 PRELIMINARY; PRT; 1881 AA.
AC Q8RGK2; DT 01-JUN-2002 (TremBrel. 21, Created)
DT 01-JUN-2002 (TremBrel. 21, Last sequence update)
DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
DE Hemolysin.
GN FN0291.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyripides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AEO10541; AAL94497.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006838; Haemagg.act.
DR InterPro; IPR001452; SH3.
DR Pfam; PF05860; Haemagg.act; 1.
DR Pfam; PF00018; SH3; 1.
DR PROSITE; PS00761; SPASE_I_3; 1.
KW Complete proteome.
SQ SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;
Query Match 6.0%; Score 226.5; DB 16; Length 1881;
Best Local Similarity 20.9%; Pred. No. 0.026;
Matches 176; Conservative 131; Mismatches 302; Indels 233; Gaps 37;
Qy 30 QAPNVTS-----STTG-----DLSIPSSLE---NIPSENYQ-FQSAIWSGPI 69
Db 246 QAPNYATGDVIVSSKGVKVLKDTQAKRDIKISSTETETIGSKLLAENAINISGKTSNG 305
Qy 70 KVKPSDEYFATGADNHVTMVDDQEVIN-----KASNSNKIELEKGR 112
Db 306 QIRANNITINGNVDSNLFTKDTISGNLKNNGSVSSNLNVKEIENSKNVVVE-K 364
Qy 113 LYVQIKI-----QYQENPTEKGLDKLWY--DSQNKKEVSSDNLQIPELKQKSS 161
Db 365 LSSTKITNLGNLSAKIEKTNFNSGLFKSKNITAKDFKNNGEV-SSENLTNTNLENSNK 423
Qy 162 -----NSRKRSTSGAGTPVPRDNDG----- 182
Db 424 INVKENINSINYKNTAETTSKNTLNLDNRGNTIINNVSSGVIANNGKLLVGNITIN 483

Qy 183 ---IPDSLEVEGTVDVKNKRTFELSPWISNIHEKKGLTKYKSPFKWSTASDPYSDFEKV 239
Db 484 SONLTNTATVOGKTDFDKNKINSNGKILSDNLTKOIFSSGNISAKVITTOELINGEII 543
Qy 240 TGRDKNVSPEARHPLVAAAYPIVHVDMEN-----IILSKNEDQSTQNTDSETRTISKNT 293
Db 544 SNNLSSNNNNKXIFVNGNLKISNNLANSVGIEGLELNTNSIENTGNIITKNKLSQNL 603
Qy 294 STSRHTSEVHGNAEVASFDDIGGSVAGFSNSNSTVA---IDHLSLAGERTWAETM 350
Db 604 NNKK-----NTANVAGFLDVHKNISS-VGNIKAITMTKNTNLDNSGNIL-----TN 648
Qy 351 GLNTAD-----TARLNANIRVNTGTAFIYNVLPPTSLVLGKNQTLATIDADENQLSQ 403
Db 549 SLTTAENINKSITAKNISQNLVNSGS-----VISDNITVA-----DN 687
Qy 404 ILAPNNYPSKNLAPIALNAQKPSPTITNNYNQFLELEKTKQLRLDTPQVGNINATYN 463
Db 688 ITNTNNIF-----ANEKISADKIS-NSNK-LVAKNTEITTKLTND---GNIVVK- 730
Qy 464 FENGRRVVDGTGSNWSEV-----LPQICETARIIFNGKDLNLVERRIAAVNPSPDPLE 515
Db 731 -ENLKAKDITNSNTIKVGENLNTDKLQNSKTLAKNINIEKSLNNGKITSLNANINTS 789
Qy 516 TTKPDMTLKALKIAFGNPNPGLNQYQGXDIETFFDNFPOQTSQNIKNQALAEINATNIY 575
Db 790 DIKNNGIIQAIK-----NINIKTSNDLKLKGKYTANDSLNINAKSLENNGN--- 836
Qy 576 TVLD---KIKLNKAKMNTILIRKRFHYDRNNAIV-----GADESUVVKEARREVINS--T 624
Db 837 --LENDGKIRFNLGTNNLNKISSSNLNTANITANEISNNGVNSIIIGSEAKLTITANSLKN 894
Qy 625 EGGLE-----NI-DKDIRKILSYVIEIDTEGLKEVINDRYDMLNISSLRQ 670
Db 895 EGNLFPGEIENKLTGNTTNTGTVISSLGLKIEAKDVNDKXIIISDNDLTDIVNSITN 954
Qy 671 DGKTF-----IDPKK--YNDKLPILVSPNKNVYVYTKENIINPSENGDSTNGIK 722
Db 955 KGLLYSTNNMKVDPKFNFLNDKAEIVSSG-----DITN-SENG-TFTNRVG 999
Qy 723 KI 724
Db 1000 DI 1001
RESULT 11
Q8II04 PRELIMINARY; PRT; 3468 AA.
AC Q8II04; DT 01-MAR-2003 (TremBrel. 23, Created)
DT 01-MAR-2003 (TremBrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF11_0371.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Petrea M., Allen J., Seiwung J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).

```
QY 226 WSTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEQDQSTQNTDSE 285
DB 231 PCTANDPYDPEKVSQGDPSVMSVARDPMISAYPIVGVQMERLVVSKSE-----TITGDS 286
QY 286 TRTISKNTSTGRTTSEVHGNAEVHASFDDIGG-----SVSAGFSNSNSTVAIDHSL 338
DB 287 TKSMKSTSHSTNINTV--GRAVSGSLQAGGIPPVFMSASANSYSHTWQNTSTVDDT- 343
QY 339 SLAGERTWAETWGLNTADTARLANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIDAE 398
DB 344 --TGE-SFSGSLINTGSSAVINPIRYNTGTAPVYVNTPTTIVIDK-OSVATIKGOE 399
QY 399 NOLSQILAPNNYPYPSKNLAPIALNAQKFFSTPTIMYVNFQLEKTKQLELDTDOQVYN 458
DB 400 SLIGDYLNPGGTYPIIGEPMALNTMDOFSSRLPIYVNLQKSIDNGGTWMLSTSQFTGN 459
QY 459 ZATYFENGRVAVDTGNSWSEVLPOIQETTRAIL--FNGKDLNVERRIAAVNPSPDLET 516
DB 460 FAKYN-SGNLVTG--GNWGPYLGITKTTASLTLSFSGQTQVA---VWAPNPSDPEDK 514
QY 517 TKPDMTLKALKIAPG--FNEPNGNLOQYQCKDIT--EPDFNFDQQTSONIKQLAELNATN 573
DB 515 T-PKULTSQALVKALFKKQKGFYFHGLEISKNEKIQVFLDSNTNDFENQLKNTADKD 573
QY 574 IYTVLDKIKLNAAQNMILIRDRPHYDRNNIAV 605
DB 574 IMHCI--IKRN--MMILVKVITFKENISSINI 601
RESULT 8
ID Q844J8 PRELIMINARY; PRT; 775 AA.
AC Q844J8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ViplAC.
GN ViplAC.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y., Chen J., Pang Y.;
RT "Cloning of vipA(C) and vip2A(C) from Bacillus thuringiensis HD201.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY245547; AAC86514.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; BINARY_TOXINB.
DR PRINTS; PR01391; BINARY_TOXINB.
SQ SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;
Query Match 19.1%; Score 721; DB 2; Length 775;
Best Local Similarity 29.0%; Pred. No. 5.2e-26;
Matches 223; Conservative 133; Mismatches 294; Indels 120; Gaps 30;
QY 4 QENRLNESSSQGLLYGYPDLNFQAPMVVTSSTTGSLSPSSLEN--IPSENQYFQ 61
DB 42 QKQKQ--QKEMDRKGLLYGFKGDF-SNLTWPAFTRDNTLIYDQQTANKLDRKQYQ 97
QY 62 SATWSGFIKVKSDVEYFAFSADNHVTWVDDEVINKASNSKIRLEKGRLYQIKIQY 121
DB 98 SIRWIGLQSKGKDFNFNSEBQALIEDGKLIISNKGKQVHLEKLVPIKIEYQ 157
QY 122 RENPTEKGLD-----FKLWYTSQNKKEVISSDNLQLPELKQKSN----- 162
DB 158 SD--TKFNIDSKTFKFEKFLKIDSONSQOVKRDLENFENFKKESREFLAKASKTNFM 215
QY 163 SRKKRSTSGAFTVDRDNDGIPDSLEVEGTVDVKNKRTFLSPWISNIHEKGLTKYKS 222
```

```
DB 216 QKMKRDIDED--TDTGDSIPDLWEENGYT--IQNK--VAVKW--DKFAQQQYKYKLSS 267
QY 223 PEKWTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEQDQSTQNT 282
DB 268 PYQAHITVGDPYDWEKAAGDIPKSNAAATRPLVAAFPSINVDMRKMLSKDSNLS----- 323
QY 283 DSETRITSKNTSTGRTTSEVHGNAEVHASFDDIGGVSAGFSNSNS--TVAIDHSL 339
DB 324 -----NSAEAHNSNSYTVANSE-----GASIEAGFGPKGFGVGSANYQHTET 366
QY 340 LAGERTW-----AETWGLNTADTARLANIRVNTGTAPIYVNLPTTSLVLGKQNTLATID 395
DB 367 VGSD--WGNKSENTQFNSASAGYLNANVHNVNVTGGIYDQPTTSFIL-QDSFIATIT 423
QY 396 ADENQLSQILAPNNYPYPSKNLAPIALNAQKFFSTPTIMYVNFQLEKTKQLELDTDOV 455
DB 424 AKSNATALSIPSGDRYPASK-EGISLKTWDDDNSHPITLKNPQLDAVLNNEVIKINTDOT 482
QY 456 YGNATYFENGRVAVDTGNSWSEVLPOIQETTRAILFNGKDLNVERRIAAVNPSPDPLE 515
DB 483 DCRYGIIGVDG--KAEIGDRWSPIDIEIKGTASIIIDFADGKALETRIAAKDYKNPED 539
QY 516 TTKPDMTLKALKIAPG--FNEPNGNLOQYQCKDIT--ITEFDN-----FDQQTSON 561
DB 540 KT-PSLTIKEGKLIAPESISEDKGILFYEYKNDGKVTKKQLSEENIMPYLDEDTSK 598
QY 562 IKQLAELNATNIYTVLDKIKLNAAQNMILIRDRPHYDRNNIAVGADESIVKEAHREVIN 621
DB 599 FERQLSDGSAKGLY----DKLTPKKNITIR-----LATVTLGDFDQFSAYPWENATW 647
QY 622 SSTEGLL-----LNIDKDIRK-----ILSGYIVE--IEDTEGLKEVI---NDR 659
DB 648 SDFGNLRGLSLAIPQESKYTPKQVKNPYDYLITGYIKHDFTTDNESLGI-VATKKDN 707
QY 660 YDMLNLSLRQDGKTFIDPKYNDKLPYISNPNKYVAVYTKENTIN 709
DB 708 FEMWNGTSIFSONSGGEFKFTIKTQ-NISG-DYILDSIQLMKRNDVN 755
RESULT 9
ID Q8KYK2 PRELIMINARY; PRT; 225 AA.
AC Q8KYK2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protective antigen-related protein, (pXOI-111).
GN BXA0163.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2012;
RX MEDLINE=22061436; PubMed=12004073;
RA Read T.D., Salzberg S.L., Pop M., Shumway M., Umayam L., Jiang L.,
RA Holtzapflee E., Busch J.D., Smith K.L., Schupp J.M., Solomon D.,
RA Keim P., Fraser C.M.;
RT "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in
RL Science 296:2028-2033(2002).
DR EMBL; AB011190; AAM26108.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 225 AA; 25402 MW; 2E121BE54295F9C8 CRC64;
Query Match 7.9%; Score 296.5; DB 2; Length 225;
Best Local Similarity 37.2%; Pred. No. 8.9e-07;
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;
QY 566 LAELNATNIYTVLDKIKLNAAQNMILIRDRPHYDRNNIAVGADESIVKEAHREVINSTE 625
```

RN SEQUENCE FROM N.A.
RC STRAIN=NCIB 10748;
RX MEDLINE=84041637; PubMed=8225592;
RA Ferelle S., Gilbert M., Boquet P., Popoff M.R.;
RT "Characterization of Clostridium perfringens iota toxin genes and
RL expression in Escherichia coli."; Infect. Immun. 61:5147-5156(1993).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=NCIB 10748;
RA Popoff M.R.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X73562; CAA51960.1; -;
DR FIR; I40862; I40862.
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; F:toxin activity; IEA.
DR Pfam; PF03495; B:binary toxin; IEA.
DR PRINTS; PRO1391; BINARYTOXINB.
FW SIGNAL.
FT SIGNAL.
FT CHAIN
SQ SEQUENCE 212 875 IOTA TOXIN COMPONENT 1B.
Query Match 23.1%; Score 871; DB 2; Length 875;
Best Local Similarity 31.1%; Pred. No. 5.5e-33;
Matches 252; Conservative 131; Mismatches 278; Indels 150; Gaps 31;
QY 1 EVKQENRLNESSESSGGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSE-NQY 59
DB 35 DTNKEETINENTLSSNGLMGYFDEHFKDLELAPIKNGDLKFEKVKDKLLTSS 94
QY 60 FQSAIMSGPIKVKSDYEYFATSADNHVTMVDQEVINKASNKIRLEKGRLYQIKI 119
DB 95 IKSIRWTGRIIPSEDEYILSTDR-NDVLQINAKGDIK---TLKVNMGKQAVNRIE 150
QY 120 YQEN-----PTEKGLDKLYWDSQNKKEVISSDNLQLPKQKSSNRKRTS 170
DB 151 IQDKNIGSDNLVSP-----KLYW-ELNGKNTVPEENLFFRYSKIDEND----- 195
QY 171 AGPTVP-----DRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIH 211
DB 196 --PFIFNNFFDVRFPSSAAWEDLTDNDNIPDAVEKNGYI---KDSIAVKNDQSA 249
QY 212 EKGLTKYSSPEKWTASDPSDFEKTGRIDKNVSPPEARPLVAAYPIVHVDMENIL 271
DB 250 E-QGYKKYSSVLESNTAGDPTDYQKASGSDKAIKLEARDPLVAAYPVVGVMENLII 308
QY 272 SKNEDOSTQNTDSERTIGKNTSTSTHTSEVHGAEVHASFDDIGGSVSAFNSNGST 331
DB 309 STNEHASS-----DQKTVSRATNKTNDANTV-----GVISAGYQNGFTGN 351
QY 332 VAIDHS-----LSLAGERTWAETMGLNTADTARLNANIVNTGTAPIYVNLPTSL 383
DB 352 ITTSYGHNTDSTAVQDSNGESWNTGLSINKGESAVINANVRYNTGTAPMYKVTPTNL 411
QY 384 VLKKNOTLATIDENQLSQILAPNNYPSKNLAPIALNAOKFSTPTTWNQVLFLE 443
DB 412 VL-DGELTATIAQDQGNUSPNETPKGLSPALMTQFQARLPIYNDQKLLD 470
QY 444 KTQLRLDQVYGNATYFENGVRVDTGNSWSEVLPQIQETTARIIFNGKOLNLYER 503
DB 471 SGQIKLETTQVSGNYGTKN-SQGI-I-TEGNSWSNYISQIDSVSASIIID-TGSOTFER 527
QY 504 RIAAVNPSPLETKPDMILKALAFGNFENGNYQY-QGKDITE--FDNFQOQTSQ 560
DB 528 RVAAKSGQNPEDKT-PEITIGEAIKKAFSATK-NGELLYFNGIPIDESCVLEIFDNTSE 585
QY 561 NIKNQLAELNATNIYTVLQIKLNKONLIIIRDKRF--HYDR-NNIAVGADESVVKEAHR 617

586 I IKEQLKYLDKKIYV---KLRGMNLIKVPSTYTFNDFEYNNFP--ASWSNIDTKNQ 639
QY 618 EVINSTEGL-----LLNIDKDIRKLSGY-----IVEIEDTEGLKEVIND 658
DB 640 DGQSVANKUSGETKIIPMSKUKPKRYVFSGYKDPSTNSITVNIKSKEOKTDYLPV 699
QY 659 RYDMLNIS-----SLRQDGKTFIDFKKYND--KLPLYISNPYKV----- 696
DB 700 EKDYTFSEYFETTKGDSSDIETLTATSSGVFLDNLSTELNSTPEILKEPEIKVPSDQE 759
QY 697 -----NVYATVTKENTIIINPSENGDTSTNGI 721
DB 760 ILDAHKKYADIKLDT-----NTGNTYIDGI 785
RESULT 7
ID 086171 PRELIMINARY; PRT; 721 AA.
AC 086171
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE C2 toxin (Component-II).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=98323874; PubMed=9659689;
RA Kimura K., Kubota T., Ohishi I., Isogai H., Isogai E., Fujii N.;
RT "The gene for component-II of botulinum C2 toxin.";
RL vet. Microbiol. 62:27-34(1998).
[2]
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=96184657; PubMed=8645309;
RA Fujii N., Kubota T., Shirakawa S., Kimura K., Ohishi I., Moriishi K.,
RA Isogai E., Isogai H.;
RT "Characterization of component-I gene of botulinum C2 toxin and PCR
RT detection of its gene in clostridial species."; Biochem. Biophys. Res. Commun. 220:353-359(1996).
RL EMBL; D88982; BAA32537.1; -;
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB.
DR PRINTS; PRO1391; BINARYTOXINB.
SQ SEQUENCE 721 AA; 80515 MW; 44C8153AC749D5F2 CRC64;
Query Match 21.5%; Score 813; DB 2; Length 721;
Best Local Similarity 35.3%; Pred. No. 2.3e-30;
Matches 223; Conservative 118; Mismatches 227; Indels 64; Gaps 25;
QY 3 KQENRLNESSESSGGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
DB 5 KFENSVNSKNQYFTINGLMGYFEN-DFENLIIISPTLDNLTFSKEDINSILG-NKII 62
QY 61 QSAIWSGFIKVKSDYEYFATSADN-HVTMVDQEVINKASN-SNKIRLEKGRLYQIKI 118
DB 63 KSAEWIGLIRPSITGEYILSTNSPNCRVEL---NGEIFNLSTNTNTVNIQGNVYDRI 119
QY 119 -QYQRENPTEKGLD-PKLYWTDQSNKKEVISSDNLQLPKQKSSNRKRTSAGPTV- 175
DB 120 EQLMSENQLKKNYEGIKLYWETSIDIKEIIPSEVL-----LKPYSNTNEKSKTIPNLT 175
QY 176 -----PDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEK 225
DB 176 SNALKKANANRDTDRDGPDEWEINGYTVNMQKAVAWMDKFAAN-----GYKKYVSNPFFK 230

```
Db 321 ----DOGKTVSRNTTNSKTDANTAGVAINIAVQNGF---TSGITTYNSHTTENSTAVQNS 373
Qy 338 LSIAGERTWAETMGLTADTARLANANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDAD 397
Db 374 ---SWNTSLSINKGESAYINANRYVNTGTAPYKVTPTTNLVL-DGDTLTITIKAQ 428
Qy 398 ENOLSOILAPNNYPKSNLAPIALNAQKFSPTPTMNYNQFLELEKTKQLRLDQVYVG 457
Db 429 DNOIGNNLSENYPKKGLSPLALNTWDFSRLLPINTYDQKKJAGKQIKLSTTQVSG 488
Qy 458 NZATYFNGRVRVDT-GSNWSEVLPOIQTETARIIFN-GKOLNLVERRIAAVNPSDPLE 515
Db 489 N---YGIKNSQGGIITEGNSWSYISQIDSLGASIIDTGSQ--VFERSVTAKDSNPED 543
Qy 516 TTKPDMTLKEALKIARFNEPNCNLOVQKGDITE--FDNFQOQTSQNIKNQGLASINATN 573
Db 544 KT-PVITIGAEIKAFKATKNGILYFNGMPIDESCVELIFDGTANTLIRERLNALNDKK 602
Qy 574 IYTVLDKIKLNAKXNLIIRD-----IAVGADSVVKEAHEV-----I 620
Db 603 IYNV---QLERGKILIKTSTYFNNFDGYNPPSSWSNVDSNNQDGLQNAANKLSGETK 658
Qy 594 -----KRFHY----- 598
Db 659 IYPMKLNPKRYVFSGYLKNSTSNPIYVNIKAEQKTYNLVSNFYKFSYEFETIG 718
Qy 599 -DRNN-----IAVGADSVVKEAHEV-----I 620
Db 719 RDASNIEITSSGTFLDNLSTELNSTPEILKEPDIKVPSDQBI-DHKKYADLSF 777
Qy 621 NSSTEGLLN-----IDKDIRKILSGVIVIE-DTEGLKEVINDRYDMLNISLROD 671
Db 778 NQSTANYLDGLYFEPTQTNKEVDYIQYKVEATLEYSGFKDIGTKKELANYTGDSNQ 837
Qy 672 GKT-FIDFKYNDKPLUYISNPY---KVVYAVTKNTIINPSENGDTSTNGIKKILI 726
Db 838 PKTYNVNFRSY-----FTSGENVMPYKLIIRIATPEN-----KELLV 875
Qy 727 FS 728
Db 876 LS 877

RESULT 5
Q9KH41 ID Q9KH41 PRELIMINARY; PRT; 876 AA.
AC Q9KH41;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CdB.
GN CDB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
CX NCBI_TaxID=1496;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 20309;
RA Chang S.Y., Song K.P.;
RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain
RL submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF271719; AAF81761.1; -.
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; F:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;
```

```
Query Match 23.2%; Score 875.5; DB 2; Length 876;
Best Local Similarity 31.4%; Pred. No. 3.4e-330;
Matches 250; Conservative 140; Mismatches 280; Indels 125; Gaps 30;
Qy 4 QENRLNESSSSQGLGYFSDLNFPQAPMVVTSSTTGLSIPSSLEN-IPSENGYFOS 62
Db 40 KKKEIVNEDILPNNGLMGVYFTDEHFKDLKMAPIKDGNLKPEEKVLDLDDKSDVKS 99
Qy 63 AIWSGFTKVKKSEDEYTPATSDAHNVHTWVDDQEVINKASNSKIRLEKGLVQIKYQOR 122
Db 100 IRWTGRIPSKDGEYTLSTDRDD-VLMQVNTTESTI---SNTLKNVKKKGEYKVRLELQD 155
Qy 123 EN--PTEKGLDFKLXYTWDSONKKEVSSDNLQLPKQKSSNRKRSSTAGTVP--- 176
Db 156 KNLGSIDNLSPNLYW-ELDMKKIIPENLFLRDY---SNIEKD-----DFIFPNNNF 205
Qy 177 -----DRNDGIPDSLEVEGTVDKVKRTPFLSPWISNIHKKGLTKYKSS 222
Db 206 FDPKLMGDEWEDLDTDNDNIPDSYERNGYTI----KDLIAVKEWDSFAB-QGYKKYVSN 260
Qy 223 PEKWTASDPSYDPEKVTGRIDKNVSEARHPLVAAVPIVHVDMENIILSKNEDQSTONT 282
Db 261 YLENTAGDPTDYEKASGDFDKAIKTEARDPLVAAVPIVGVGMEKLIISTNEHAST-- 317
Qy 283 DSETRTISKNTSTSRHTSEVHGNAEVHASFIDIGSVSAGFSNSNSSTVAIDHS----- 337
Db 318 -DOGKTVSRATTSKTESNTA-----GVSNVGYQNGFTANTVYSHTTDN 363
Qy 338 ---LSLAGEITWAEWTGLNTADTARLANANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATI 394
Db 364 STAVQDNGESWNTGLSINKGESAYINANRYVNTGTAPYKVTPTTNLVL-DGDTLSTI 422
Qy 395 DADENQSQILAPNNYPKSNLAPIALNAQKFSSTPTMNYNQFLELEKTKQLRLDQD 454
Db 423 KAQENQIGNNLSPGDTYPKKGLSPLALNTWDFSSRLIPINYLQKLDAGKQIKLETTQ 482
Qy 455 VYGNIAIYFNGRVRVDTGSMSEVLPOIQTETARIIFNGKOLNLVERIAAVNPSDPL 514
Db 483 VSGNFGTKN-SSGOI-VTEGNSWSYISQIDSLGASIIDTEN-ESVERVTAKNLQDPE 539
Qy 515 ETTKPDMTLKEALKIARFNEPNCNLOVQKGDITE--FDNFQOQTSQNIKNQGLASINAT 572
Db 540 DKT-PELTIGEAIEKAFKATKDGGLLYFNDIPIDESCVELIFDNTANKIKSLKLSDK 598
Qy 573 NIYTVLDKIKLNAKXNLIIRDKRFHY---DRNNIAVGADSVVKEAHEVINSTEGL-- 627
Db 599 KIYV---KLERGMNLIKPTTYFTNFDDYNNYP---STWSNVNTTKOGLQGSANKLNG 652
Qy 628 -----LLAIDKDIRKILSGY-----IVEI-----EDTEGL-----KEVI 656
Db 653 ETKIKIPMSELKPYKRVVFSGYSKDPLTSNSIIVKIKAKEKTDYLVPEQGYTKFSYEFE 712
Qy 657 NDRYDMLNIS-SLRQDKTIDFKKYND--KLPLYISNPYK-----VNVYAVTKENT 706
Db 713 TTEKSSNIEITLIGSGTGYLDNLSTELNSTPEILDEPEVKIPTDQEIIDAHKIYFADL 772
Qy 707 IINPSENGDTSTNGI 721
Db 773 NFNFS-TGNTYINGM 786

RESULT 6
Q46221 ID Q46221 PRELIMINARY; PRT; 875 AA.
AC Q46221;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Iota toxin component Ib precursor.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
CX NCBI_TaxID=1502;
```



```

AC Q32739;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE ADP-ribosyltransferase.
GN CDTB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CD196;
RX MEDLINE=97230316; PubMed=9119480;
RA Perelle S., Gilbert M., Bourlioux P., Corthier G., Popoff M.R.;
RT "production of a complete binary toxin (actin-specific ADP-
RT ribosyltransferase) by Clostridium difficile CD196."
RL Infect. Immun. 65:1402-1407(1997).
DR EMBL; L76081; AAB67305.1; -.
DR HSSP; P13423; 1ACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxins.
DR Pfam; PF03495; Binary toxin; I.
DR PRINTS; PR01391; BINARYTOXINB.
KW Transferase.
SQ SEQUENCE 876 AA; 98797 MW; 25E06E2D45CE2B3B CRC64;

Query Match 23.3%; Score 877.5; DB 2; Length 876;
Best Local Similarity 31.8%; Pred. No. 2.7e-33; Indels 125; Gaps 30;
Matches 251; Conservative 139; Mismatches 280;

QY 4 QENLLNESSSQGLLYYFSDLNFCQPMVVTSSITGDLSPSSELENIPSENQ-YFQS 62
DB 40 KKEIVNEDILPNNGLMGYYFSDHFKDLKMAPIKDGNLKFEKKVKKLLDKSDVKS 99
QY 63 AIWSGFTKVKSDYEYFATSADNHVTWYDDQEVINKASNNKIRLEKGLYQIKIQYR 122
DB 100 IRMTGRIIPSKDGEYTLSTDRDD-VLMQVNTSTI---SNTLKNMKGKYEKVRLEQD 155
QY 123 EN--PTEKGLDFLYKLTSDSNKKEVISSDNLOLPELKQSSNSRKRSTSAQTPV- 176
DB 156 KNLGSDINLSSPNLYW-ELDGMKKIPEENLFLRDY----SNIEKD----DFFIPNNF 205
QY 177 -----DRNDGIPPSLEVEGYTVDVKNKRTFLSFWSNIHEKGLTKYSS 222
DB 206 FDPKLMSDWEDELDTDNDNIPDSYERNGYTI----KDLIAVKWEDSPAE-QYKKYVSN 260
QY 223 PEKWTASDPYDFEKTGRIDKNVSPEARHPLVAAPYIVHYDMENIILSKNEDOSTQT 282
DB 261 YLSNTAGDPTDYEKASGSFKAITEARDPLVAAPYIVGVGMKLLISTNEHAST--- 317
QY 283 DSERTISKNTSTSRTHTEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHS----- 337
DB 318 -DQKTVSRATTNSKTESNTA-----GVSVMVQNGPTANVTNYSHTTDN 363
QY 338 ---LSLAGETWAEATGLTADTARLNANIRVYNTGTAPIYVLPPTSILVGNKNTLATI 394
DB 364 STAVQDSGBSWNTGLSINKGESAYINAVRYNTGTAPWKVPTPTNLVL-DGDTLSTI 422
QY 395 DADENQLSQIALFNYPYKNAPIALNAQKFSSTPITMNYNQFLELEKTKQLRLDTQ 454
DB 423 KAQENQIGNLSPDGYPKGLSPLALNTWDQFSRLIPINYDQLKLDAGQIKLETTQ 482
QY 455 VYGNATYFENGVRVDTGSNWSVLPOIQTETABIIIFNGKDLNLVERRIAAVNPDL 514
DB 483 VSGNFGTKN--SSQGI-VTEGNSNSDYISIDSISIIIDTEN--ESYERRVAKNLQDDE 539
QY 515 ETTKPDMLKEALKAFGFPNEPKNLQYQKQITE--FDNFQQQTSQNIKNQLAEALNAT 572
DB 540 DKT-PELTICEATEKAFATKDGLLFYFNDIPIDESCVELLFPDNTANKIKDSLTKLSDK 598

```

```

QY 573 NIYVLDKIKLNKONILIRDKRFHY---DRNNIAGADESVVKEAHREVINSTEGL-- 627
DB 599 KIYV-----KLERGMNLIKITPYFTNFDDYNNYP--STWSNVNTNQDGLQGSANKLNG 652
QY 628 -----LLNIDKDIRKILSGY-----IVEI-----EDTEGL-----KEVI 656
DB 653 ETKIKIPMSLKPYKRVVFGYSKDPULTSIIIVKAKKEKTDYLVPEQGYTKFSYEFE 712
QY 657 NDRYDMLNIS-SLRQDCKTFIDFKYND--KLPLYISNPYKV-----NVYAVTKENT 706
DB 713 TTEKSSNIEITLIGSGTYYLDNLISITELNSTPEILDEPEVKIPTDQEIIMDAHKIYFADL 772
QY 707 IINPSENGDTSTNGI 721
DB 773 NFNPS-TGNIYINGM 786

RESULT 4
ID O06498 PRELIMINARY; PRT; 879 AA.
AC O06498;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Sb component.
OS Clostridium spiroforme.
OC Bacteria; Firmicutes; Mollicutes.
OX NCBI_TaxID=29348;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CS246;
RA Gilbert M., Perelle S., Daube G., Popoff M.R.;
RT "Clostridium spiroforme toxin genes are related to C. parvifragens iota
RT toxin genes but have a different genomic localization."
RL Syst. Appl. Microbiol. 20:337-347(1997).
DR EMBL; X97569; CAA66612.1; -.
DR HSSP; P13423; 1ACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxin; I.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 879 AA; 98739 MW; 40685ACB8E05BA01 CRC64;

Query Match 23.3%; Score 877.5; DB 2; Length 879;
Best Local Similarity 28.7%; Pred. No. 2.8e-33;
Matches 259; Conservative 147; Mismatches 259; Indels 237; Gaps 34;

QY 4 QENLLNESSSQGLLYYFSDLNFCQPMVVTSSITGDLSPSSELENIPSENQ-YFQS 62
DB 36 QEVTTREKTVSSNGLMGYYFADHFKDLMLMAPVKNKGLPEKKNVKEKLTSEKTIKS 95
QY 63 AIWSGFTKVKSDYEYFATSADNHVTWYDDQEVINKASNNKIRLEKGLYQIKIQYR 122
DB 96 IRMTGRIIPSKDGEYTLSTDKDN-VLMQINAE---GEIANTLKNVMIKGQEYSIRIEIQ- 150
QY 123 ENPTEKGLDF-----KLYWTSQNKKEVISSDNLOLPELKQKSSN----- 162
DB 151 ----DKQIGYVDDLSPPKLYW-ELNGDKTLPIPEKNLFLRDYSKIDENDPFPKDNFFDLK 205
QY 163 --SRKKESTSA-GPTVPDRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKY 219
DB 206 LKSRARLASGWDGDELDTDNDNIPDAVEKNGYTI----KDSIAVKW-EDSFAQQGYKY 260
QY 220 KSSPEKWSADPSYDFEKTGRIDKNVSPEARHPLVAAPYIVHYDMENIILSKNEDQST 279
DB 261 LSSYLESTAGDPTDYQKASGSFKAITEARDPLVAAPYIVGVGMKLLISTNEHAST 320
QY 280 QNTDSETRTTSKNTSTSRTHTEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHS 337

```

Db 20 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENISENQYF 79
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 80 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 139
QY 121 QRENTEKGLDPLKYWTDSONKKEVSSDNLOLPELKOKSSNSRKRKSTSGPTVPDRDN 180
Db 140 QRENTEKGLDPLKYWTDSONKKEVSSDNLOLPELKOKSSNSRKRKSTSGPTVPDRDN 199
QY 181 DGPIDPSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 240
Db 200 DGPIDPSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 259
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNSTSRHT 300
Db 260 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNSTSRHT 319
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARL 360
Db 320 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARL 379
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 380 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 439
QY 421 LNAQKFSSTPIITMNYNQFLEKTKQLRLDQVYGNATYVFNFGRRVVDVTSNWSSEV 480
Db 440 LNAQKFSSTPIITMNYNQFLEKTKQLRLDQVYGNATYVFNFGRRVVDVTSNWSSEV 499
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNPNGL 540
Db 500 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNPNGL 559
QY 541 QYQKDIETEDFDFNDQSTQNIKNQLAELNATNIYVLDKIKLNKMNILIRKRFHYDR 600
Db 560 QYQKDIETEDFDFNDQSTQNIKNQLAELNATNIYVLDKIKLNKMNILIRKRFHYDR 619
QY 601 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIETEDTGLKEVINDRY 660
Db 620 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIETEDTGLKEVINDRY 679
QY 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Db 680 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 739
QY 721 IK 722
Db 740 IK 741

RESULT 2

Q937W3
ID Q937W3 PRELIMINARY; PRT; 743 AA.
AC Q937W3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE pag protein (Fragment).
GN PAG.
OS Bacillus anthracis.
OG Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Carbosap;
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
RA Fasanella A., Francia M., Ciuchini F.;
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus anthracis vaccine strain 'Carbosap'";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ413936; CAC93934.1; -.
DR GO: 0005576; C:extracellular; IEA.
DR GO: 0046821; C:extrachromosomal DNA; IEA.
DR GO: 0015405; F:toxin activity; IEA.
DR GO: 0009400; P:pathogenesis; IEA.
DR InterPro: IPR003896; Anthrax toxinB.
DR Pfam: PF03495; Binary toxB; I.
DR PRINTS: PRO1391; BINARYTOXINB.
KW Plasmid.
KW NON_TER.
FT NON_TER 1 743
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5DE5A6 CRC64;

Query Match 97.4%; Score 3676; DB 2; Length 743;
Best Local Similarity 99.2%; Pred. No. 1e-164;
Matches 716; Conservative 1; Mismatches 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENISENQYF 60
Db 22 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENISENQYF 81
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 82 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 141
QY 121 QRENTEKGLDPLKYWTDSONKKEVSSDNLOLPELKOKSSNSRKRKSTSGPTVPDRDN 180
Db 142 QRENTEKGLDPLKYWTDSONKKEVSSDNLOLPELKOKSSNSRKRKSTSGPTVPDRDN 201
QY 181 DGPIDPSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 240
Db 202 DGPIDPSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 261
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNSTSRHT 300
Db 262 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNSTSRHT 321
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARL 360
Db 322 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARL 381
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 382 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 441
QY 421 LNAQKFSSTPIITMNYNQFLEKTKQLRLDQVYGNATYVFNFGRRVVDVTSNWSSEV 480
Db 442 LNAQKFSSTPIITMNYNQFLEKTKQLRLDQVYGNATYVFNFGRRVVDVTSNWSSEV 501
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNPNGL 540
Db 502 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNPNGL 561
QY 541 QYQKDIETEDFDFNDQSTQNIKNQLAELNATNIYVLDKIKLNKMNILIRKRFHYDR 600
Db 562 QYQKDIETEDFDFNDQSTQNIKNQLAELNATNIYVLDKIKLNKMNILIRKRFHYDR 621
QY 601 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIETEDTGLKEVINDRY 660
Db 622 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIETEDTGLKEVINDRY 681
QY 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Db 682 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 741
QY 721 IK 722
Db 742 IK 743

RESULT 3

O32739
ID O32739 PRELIMINARY; PRT; 876 AA.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 34.6261 Seconds
(without alignments)
6697.424 Million cell updates/sec

Title: US-09-848-909A-11
Perfect score: 3774
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKILFSSKGYBIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3676	97.4	741	2 Q937W2	Q937W2 bacillus an
2	3676	97.4	743	2 Q937W3	Q937W3 bacillus an
3	877.5	23.3	876	2 Q32739	Q32739 clostridium
4	877.5	23.3	879	2 Q06498	Q06498 clostridium
5	875.5	23.2	876	2 Q9KH41	Q9KH41 clostridium
6	871	23.1	875	2 Q46221	Q46221 clostridium
7	813	21.5	721	2 Q86171	Q86171 clostridium
8	721	19.1	775	2 Q844J8	Q844J8 bacillus th
9	296.5	7.9	225	2 Q8KYK2	Q8KYK2 bacillus an
10	226.5	6.0	1881	16 Q8GK2	Q8GK2 fusobacteri
11	209.5	5.6	3488	5 Q81T04	Q81T04 plasmodium
12	202	5.4	4688	16 Q9PQ08	Q9PQ08 ureaplasma
13	201.5	5.3	2849	5 Q81HY4	Q81HY4 plasmodium
14	200.5	5.3	1904	5 Q81KS2	Q81KS2 plasmodium
15	196.5	5.2	2940	5 Q81HP9	Q81HP9 plasmodium
16	194	5.1	1404	2 Q8RJN9	Q8RJN9 mycoplasma

ALIGNMENTS

RESULT 1

Q937W2 PRELIMINARY; PRT; 741 AA.
ID Q937W2
AC Q937W2; (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Pag protein (fragment).
GN PAG.
OS Bacillus anthracis.
OG Plasmid pXO1.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ferrara;
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M., Paganella A., Francia M., Ciuchini P.
RT "Sequence analysis of the genes encoding for the major virulence factors of Bacillus anthracis vaccine strain 'Carbosap'";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ413937; CAC93935.1;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax_coxinB.
DR Pfam; PF03495; Binary_coxB; 1.
DR PRINTS; PR01391; BINARYTOXINB.
FT NON_TER 1 741
FT NON_TER 741 741
SQ SEQUENCE 741 AA; 83153 MW; C7F95820B73065C0 CRC64;

Query Match 97.4%; Score 3676; DB 2; Length 741;
Best Local Similarity 99.2%; Pred.No. 1e-164;
Matches 716; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 EVKQENRLNSESQGLLYFFDLNFQAPMVVTSSTGDLSPSELENIPSENYQF 60

Q897H6 clostridium
Q98PM9 mycoplasma
Q81L45 plasmodium
Q81548 plasmodium
Q81FP4 plasmodium
Q814R2 plasmodium
Q8NWQ6 staphylococ
Q8WSK5 plasmodium
Q26216 plasmodium
Q7YWE7 plasmodium
Q7YWE6 plasmodium
Q81G0 plasmodium
Q81GX1 bacillus ce
Q25579 helicobacte
Q815A6 plasmodium
Q9GZ76 plasmodium
Q7YWF0 plasmodium
Q8147 plasmodium
Q9BJX9 plasmodium
Q9AV25 oryza sativ
Q7XC10 oryza sativ
Q81C27 plasmodium
Q813R5 plasmodium
Q99U54 staphylococ
Q931R6 staphylococ
Q81H2 plasmodium
Q8118 plasmodium
Q81B09 plasmodium
Q81531 plasmodium

Db 537 LKGEIQTEISNBEHLSQSQTAAEKEAAVATNNELSESK--NSLQTL-CNA----- 585
Qy 371 TAPIYVLPPTTSILVLGKN-QTLATIDADENQJLS---QILAPNNYYPSPKNLAPIALNAOKK 426
Db 586 ---FQEKLAQSVYQLKENEQNFSSLDTSFKLNESHQBLENNHQITTKQL-----KD 634
Qy 427 FSTPTITMYNOFLELEKTKQLRLDQDVYGNIAFYNFENGVRVDTGNSWSEVLPIQOE 486
Db 635 TSS-----KLQQLERANFEQKESTLSDENNDLRTK-----LLKLEE 672
Qy 487 TTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDW-TLKEALKI----AFGFNEPENGNIQ 541
Db 673 SNKSLIKKQEDVDSLEKNI-----QTLKEDLRKSEEARFSKLEAKNLREVIDNL- 722
Qy 542 YQCKDIT-BFDNPDQOOTSQNTKNQLAELNATNIYTVLDKIKLNAMMILIRDRPHYDR 600
Db 723 -KGKHETLEAQRNDLHSSSLSDAKNTNAILSSBLTKSSESDVKELTANVETLTQDSK----- 776
Qy 601 NNIAGADESVVKEAHREVINSGTEGLLNIDKDIR-----KILSGVIVEIE-----DT 649
Db 777 -----AMKQSTSLVNSYQS--ISNLYHELRDDHVNMQSQNTLLESESKLTDC 824
Qy 650 EGLKE---VINDRYDMLNSSLRQDGKTFIDPKKYNKDLPLYISNPNYKVNVAVTKENT 706
Db 825 ENLTQQNMTLIDNVQKLMHKHVNQESKV-SELKEVNGKLSLDLKNLRSLLAV-AISDNDQ 862
Qy 707 IINP-----SENGDT-----STNGIKKI 724
Db 883 ILTQLAELSKNYDSLEQESAQLNSGLKSL 911

Search completed: May 3, 2004, 19:36:59
Job time : 12.2539 secs

Db 843 SVKNGVNGTLV-----GNLSKARATT-----LSKNFSDIKKELNAKLG-----NFNN--- 885
 QY 442 LEKTKQLRLDTQVYGNATYFNENGRVAVDTGSGNWSVLPQIQETTTARIIPNGKDLNLA 501
 Db 886 -NNNNGLENSTEPIYQVA-----KKYKAKIDRLDQIASGLGVDGQAASFLKKRHKVDL 940
 QY 502 ERIIAVNPSPDLETTKPM-----TLKEALKIAFGFENEGNQLYQGDITFDFNFQOQ 557
 Db 941 SKVGLSAN-HEPIYATIDLDGPPFLKRDKV-----DLSKVGLSRBOK 984
 QY 558 TSONIKN---QLAELNAT---NIYTVLDIKLNAKNNILIRKPHYDRNNIAVAGDSV 611
 Db 985 LTQKIDNLNQAQVSEAKASHFDNLQMDIKLKDSTKKNV-----NLYVESAKV 1033
 QY 612 VKEAHEVINSSTEG---LLNLI-DKDIRKILSGVIEVEDTEGLKEKVINDRYDMLNISS 667
 Db 1034 P7SLSAKLDNYATNSHTRINSVYKGTINEKATGMLTQ-KNSEWLK-LVNDKIVAHVGS 1091
 QY 668 LRQDGTFTDF-----KYND--KLPLYISNP-----NYKNVYAVTKENIINPS 711
 Db 1092 APLSAYDKLGFQKQKMDYSDFKSTRLSNVAVKIKSGFVQLNIFSMGYSILMKASV 1151
 QY 712 ENGDTSTNGIKILIFSCKGYE 733
 Db 1152 EHGKVNNTN-----TKGQFQ 1165

RESULT 15
 SPOF_SCHPO STANDARD; PRT: 1957 AA.
 AC Q10411; Q9USE9;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Sporulation-specific protein 15.
 GN SPO15 OR SPAC1F3.06C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 [1]
 RN SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
 RX MEDLINE=201071136; PubMed=10639340;
 RA Ikemoto S., Nakamura T., Kubo M., Shimoda C.;
 RT "S. pombe sporulation-specific coiled-coil protein Spo15p is localized
 to the spindle pole body and essential for its modification";
 RL J. Cell Sci. 113:545-554(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris B., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA Jones K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicikert G., Aert R., Robben J., Grynolprez B.,
 RA Waltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer C., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehtz H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Gaffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 RN [3]
 RP SEQUENCE OF 705-871 FROM N.A.
 RC STRAIN=968 190;
 RX MEDLINE=20233869; PubMed=10759889;
 RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
 RA Hiraoka Y.;
 RT "Large-scale screening of intracellular protein localization in living
 fission yeast cells by the use of a GFP-fusion genomic DNA library";
 RL Genes Cells 5:169-190(2000).
 CC -1- FUNCTION: Has a role in the initiation of spore membrane
 formation.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Spindle pole body.
 CC -1- SIMILARITY: Belongs to the MPC70 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z70690; CAA94624.1; --
 DR EMBL; AB027811; BAA87115.1; --
 DR PIR; T38077; T38077.
 DR GeneDB; Spombe; SPAC1F3.06C; --
 KW Sporulation; Coiled coil.
 FT DOMAIN 199 785 COILED COIL (POTENTIAL).
 FT DOMAIN 804 1235 COILED COIL (POTENTIAL).
 FT DOMAIN 1320 1471 COILED COIL (POTENTIAL).
 FT DOMAIN 1481 1723 COILED COIL (POTENTIAL).
 FT SEQUENCE 1957 AA; 222785 MW; 3F480CA0617D9DA CRC64;
 SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA0617D9DA CRC64;
 Query Match 4.2%; Score 158; DB 1; Length 1957;
 Best Local Similarity 20.3%; Pred. No. 2.7;
 Matches 176; Conservative 134; Mismatches 327; Indels 232; Gaps 41;
 QY 2 VKQENRLLNSE-----SSQGLLYGYFDLNFQ---APWVTSSTGDLSPSSE 49
 Db 129 VTQSNLNLKLVQVRSKLAALAHENGILSLQSSNKKKXNTSVTLTSEEDVSYFQKK 189
 QY 50 LENIPSENQYFQSAIWS---GFIKV-----KKSDEYTFATSDAHNVMTWYDDQEVINKAS 101
 Db 189 LTNWESNFSKQSEAYDLRSQLLTVTEKLDKKEKDYKEDVSSIKASLAEEQASNKS 248
 QY 102 NSNKIRLEK---GRLYQIKYQRENTPE---KGLDFKLYWTSQNKKEVSSDNLQIPE 155
 Db 249 RGEQERLEKLVSSNKTSTVTLRQTENSLRAECKTLQEKL-----EKCANEEDSKLEE 302
 QY 156 LKQKSSN-----SRKKRSTSGAPTVPDRNDGIPDSLEVEGYTVDVKNKR----- 200
 Db 303 LKHNVAYSDAIVHKDKLIEDLSTRIFEFN-----LKSERDTLSIKNEKLEKLLRNTI 356
 QY 201 -TFSPMISNHEKXGKTKYKSS---PEKWTASDPYSDFEK-----VTGRID----- 244
 Db 357 GSLKDSRTSKSOLBEEMVELKESNRTHISQTLDAESKLSSEFQENKSLKGSIDYQNLS 416
 QY 245 -----KNVS---PEARHPLVAA-----YPIVHVDN----- 266
 Db 417 SKDNVQVSSQLEEARSSLAHATGKLAENISERDFQNKIKKDFEKTQDRLACLNSSN 476
 QY 267 ---ENILSKNEDQSTQNTD---SETRTISKNTSTS-----RTHTSRVHGVNAEVHAFPD 315
 Db 477 ELKEKSLIDKKDQELNNLREQIKQKQVSESTQSSLSQSLQRLDILNEKKKHEVYESQLNE 536
 QY 316 IGGSVSAGFSNSN-----STVAIDHSLSLAGERTWAGTWTGLNTADTARLNANIRYNTG 370

FT CONFLICT 887 887 T -> P (IN REF. 2).
FT CONFLICT 1284 1284 Y -> S (IN REF. 2).
SQ SEQUENCE 1420 AA; 160000 MW; 7f6CF4BBE0FAC918 CRC64;

Query Match
Best Local Similarity 4.2%; Score 159; DB 1; Length 1420;
Matches 155; Conservative 129; Mismatches 287; Indels 242; Gaps 40;

QY 27 LNFQPMV-VTSTGDSIPSELEINSENOYFCSAI--WSGFIKVKSDY-----77
DB 274 LIFAQPVADIENTSGD-----PSEFHCLODALDAIDDIQLKQTAAYRTPOSS 322
QY 78 -----TFATSDNHNVTMWVDDQEVINKASNKIRLEKRLYQI-----KIYQREN 124
DB 323 GVLSSNIAGTNPLSSDGYT-----EQFQHYKN-NSISSQFASVHVSQETNKISPKDPS 375
QY 125 PTEKGLDFKLYTDSQNKKEVVISDNMLQPLPELKQKSSNRKRSAGTVPDRNDGIP 184
DB 376 PNFPTGDKMLSPSDQAFAPLTPNNNINE--NELFNDRKQTIVS-----NDLENSPLK 428
QY 185 DSLEVEGYTVVKNKRTFTSPISNIHEKKGLTKVKSPEKW-STASDPYSDFE-----237
DB 429 TELEANGRSLEKVNNSVKTGSDVDTLNKEGLEQREQENILPSDKSDSWDKELFGDE 488
QY 238 -----KVTGRIDKNVGEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSET 286
DB 489 DEDLFGDSNKSNTSNEKSISDEI-----TDMFEMSDEEENNN-----529
QY 287 RTISKNTSTSRTHTSVHGNAEVHASFDDIGGSVAGRSNSNSTVAIDHSLSLAGRTW 346
DB 530 KSIKN-----NKEWHT--DLGKDIPF-FPSSSEKPI-----RTM 561
QY 347 AETMGLNTADTARLANIRYVNTGTAPIYV-LPTSLVLGKKNQTLATIDADENQSOIL 405
DB 562 SGT-----TKLNGKRKYLD--IPIDEMTLPTSPLYMDPGAPL-PVETPRDRKRSVF 610
QY 406 APNNYYP-SKNLAPIALNAQKESSTPI-----TMNYN-QFLELEKTKQLRLTDQVYGN 459
DB 611 APLNFNFIENNVNKNYKSGGKFSFSPLOKEERALNFDISMADLSSEE--BEDDEENG 667
QY 460 ATYNFENGVR-----VDTGSNWSE-----VLPQIOETTAII-----FNG-KDILN 499
DB 668 SDEDLKSLNVRDKMPSDNISTNTNIHEPOYINYSIPSLQDS---IIKOENFNSVNDAN 724
QY 500 LVERR-----IAAVNSPDLETTKPDWTLKEALKIAFGNEPNGNLQYQKGIITFEDNF 554
DB 725 ITSKEGFSIWIKPONDIPQTESPLKTVDSII-----757
QY 555 DQTSQNIKNQLAELNATNIYTVLDIKLNAKNILIRDKRFHYDRNNIAGDSVWKE 614
DB 758 -QPIESNIMKTLNENNVTSNPSEFTPMNVSISNLPKKS-----GIPEFTPAD 806
QY 615 AHREVINSTEGLIL-----NIDKDIRKILSGYIV---EIEDT 649
DB 807 PNLSPFESSSLPFLRHMP-LASIPDIFITPTPVVTISEKEQDILDIAEQVVTYDYNILGN 866
QY 650 EGLKEV---INDRYDMLNTSSLRQDKTEIDFKYNDKLPYISNPYKNVYAVTKEN 705
DB 867 LGIPKIAYRGVKCOEGLITTLWLQFSTP-DRINGNDIISKEY---NMKQPVVYFKKH 922
QY 706 TIINPSNGD-----TSTNGIK--KILIFS 728
DB 923 ELIKVKHDSOPFIKFLNFRPPNGIKNPKSLLS 955

RESULT 14
CAGA_HELPJ
ID CAGA_HELPJ STANDARD; PRT; 1167 AA.
AC Q9ZLTI;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytotoxicity associated immunodominant antigen (120 kDa protein)

DB (CAG pathogenicity island protein 26).
GN CAGA OR CAI OR CAG26 OR JHP0495
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.J., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.";
RL Nature 357:176-180(1999).
CC -!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
CC OR FUNCTION OF THE CYTOTOXIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB001483; AAD06073.1; .
DR PIR; B71924; B71924.
DR InterPro; IPR005169; Caga.
DR InterPro; IPR004355; IVSec_caga.
DR Pfam; PF03507; Caga; 1.
DR PRINTS; PRO1553; TYPEBASSCAGA.
KW Antigen; Complete proteome.
FT DOMAIN 246 249 POLY-THR.
FT DOMAIN 882 889 POLY-ASN.
SQ SEQUENCE 1167 AA; 129729 MW; FDESE86B81CEBD0F2 CRC64;

Query Match 4.2%; Score 158.5; DB 1; Length 1167;
Best Local Similarity 19.6%; Pred. No. 1.3;
Matches 157; Conservative 145; Mismatches 313; Indels 187; Gaps 41;

QY 3 KOENRLNSESSESSQGLL---GYVFSDL-NF-----QAPMVVTSSTGDSIPSSLE- 51
DB 480 KADKALDREKTTTQGLKHDGVNFVDYNSFKYTNASKSPDKGVGAING-----VSHLEA 534
QY 52 -----NIPSENQYFQSAIWSGFIKVKSDYTFATSDNHNVTMWDD-----QEVIN 98
DB 535 GFSGVAVENFLNMLNLAITSV---VRQDLEDKLIAKGLSPOEANKLVKDFLSNKLVG 590
QY 99 KASNSKIRLE---KGRLYQIKIQYQRENTEKGLDFKLYWTDSONKKEVISSDNLOLPE 155
DB 591 KALNFNKAFAEAKNTGYDEVK---QAQKLEKSL-----KKRELEKDVAK--N 635
QY 156 LKQSSNSRKRKRSAGTVPDR-----DNDGIPDSLEVGYTVVKNKRTFTSPWISNI 210
DB 636 LESKSGNKNKMEAKSQANSQKDEIFALINKEANRDARAI-AYAQNLRKGIKRELSDKLENI 694
QY 211 HEKGLTKYKSSPEKWSASDPYSDFEK-----VTGRI-DKNVSPPEARHPLVAAYPIV 262
DB 695 N--KDLKDFKSFDEFKNGKN--KDFSKAEETLKALGSKVKDLGINFEWIS-----741
QY 263 HVDMENILSKNEDQSTQNTDSETRTISKNTSRTHTSVHGNAEVHASFDDIGGSVSA 322
DB 742 --KVENLNAALNEFKNGKNKDFSKVTQAK--SDLENSIKDVIINOKITDKVDNLNQAQSV 797
QY 323 GFSSNSNSTVAIDHSLSLAGRTWAEITMGLNTADTARLANIRYVNTG-TAPIVNLPTT 381
DB 798 AKATGDFSGV-----EQALADLNKFSKEQLAQQAQKNEDEFNFKGKSALYQ-----842
QY 382 SILVKGKNTLATIDADENQLSQILAPNNYPSKNLAPIALNAQKFSSTPTTMYNNOFLE 441

QY	104	NKIRLEK-----	GRLYQIKIQRENPTTEKGLDFKLVWTDSONKKEV	145		
Db	365	SQTNVKNRTHYSRHYLPVFANILDDNIYTIIRDGP--	NLTNKGFNIE-----	409		
QY	146	ISSDNLQL-PEIKOKSSNRKKRSAGTVPDRDNDGIPDS	LEVEGYTVDVKNKKTFLS	204		
Db	410	NSGQNIERNPALQKLSSSWDLFTKVCURLTKGNSRD-	-----DSTCIKVKQNNL---	458		
QY	205	PAISN-----	THEKKGLTKYKSSPEKSWTASDPYS-DPEKVTGRI	DNKNSPEARHPLV	256	
Db	459	PYVADKDSIQEIFENKIITD-ETNVQNY-SDKFS	LDSEILDGQVP--INFIVDPL-	511		
QY	257	AAYPIVHVDMENIILSKNE-----	DOSTQNTD-----	SETRISKNST--TERTHTSEV	303	
Db	512	-----	LPNVANPEPLNGEIEIVFYDDTIKVDYLSYYVLES	QKLSNNVENITLITSVEEA	567	
QY	304	HGNAEVHASFF-----	DICGSVSAGFSNNSSTVAJDHLSLUSAGERWASTM-	-----	350	
Db	568	LGVSNIKIYFLPSLAEKYNKGVOAGLFLWANEVVED	FTTINMKOTLDKISDVSVIIPY	621		
QY	351	---GLNADTA-RLNANIRYVNTG-----	TAPIYNVLPFTSLVIGKNQTLATID	395		
Db	628	IGPALNIGNSALRGNFNQAFATAGVAFLLGEGFPE	FTIPALGVFTFYSSIOEREKIILKIE	687		
QY	396	-----	ADENOLSOILAPNNY-----	YPSKNLAPIALNAO-----	KKFSS	429
Db	688	NCLQQRVKRWKDSYQWVSNLSRITQFNHINYQMYD	SLSYQADAIKAKIDLEYKYSYG	747		
QY	430	TPITMNYNFOLEUKTQJRLDTDQVYGNI-----	ATYNFENGVRVDTGSNWSEVL	481		
Db	748	SDXENIKSQVENIKNSLDVKI--SEAMNNINKFIRE	CSVTYLFKN-----	ML	792	
QY	482	POJOETTARIIFNGKD-LNL-----	VERRIAAVNPSPDLETTKPDMTLKEALK	528		
Db	793	PKVIDELNKFDLTKTELINLIDSHNILLVGEVDR	LKAKVNES--FENTMP-----	841		
QY	529	IAGF-FNEPENGNILOYGKIDITEFDFPDDQTS-	-----QNIKNQIAELNATNI-YTVLDKIK	582		
Db	842	--ENIFSNTKNSLL--KDIINFYFNSINDSKILS	QKNKALVDTSGYNAEVRVGDNVQ	896		
QY	583	LNAKMNTILRDKRPHYDRNNIAVCADESVVKEA	REVINSGTEGLLINIDKIRKILSGY	642		
Db	897	LN---ITYTNDFKLSSGDKIIVLNNNNILYSA--	IYENSSVSFWIKISKOLTNSHNEY	950		
QY	643	IV-ETIEDTGLKEVINDRYDMLNISLRQDGKTFD	FKYKNKDLPLY---ISNPYKVN	697		
Db	951	TTINSIEQNSGWKLCTANG---NIEWILOPVN-	-----RKYSLIFDYESLSHTGYTNK	1001		
QY	698	VYAVTKENTII	708			
Db	1002	WFFVTITNNIM	1012			

RESULT 12

YJL3 YEAST

ID	YJL3	YEAST	STANDARD;	PRT;	1803	AA.
AC	P47024;	P87192;				
DT	01-FEB-1996	(Rel. 33, Created)				
DT	30-MAY-2000	(Rel. 39, Last sequence update)				
DT	30-MAY-2000	(Rel. 39, Last annotation update)				
DE	Transposon Ty4	207.7 kda hypothetical protein.				
GN	TY4B OR YJL113W	OR J0750.				
OS	Saccharomyces cerevisiae	(Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota;	Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae;	Saccharomycetes.				
OX	NCBI_TaxId=4932;					
XP	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=S288C	/ F1679;				
RC	MEDLINE=97103775;	PubMed=9948101;				
RA	Czepluch C., Kordes E., Fujol A., Jauniaux J.-C.;					
RT	"Sequencing analysis of a 40.2 kb fragment of yeast chromosome X					

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas R., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880 (2002).
RN [2]
RP IDENTIFICATION, AND GENE NAME.
RX MEDLINE=20143585; PubMed=10679021;
RA Jin H., Amberg D.C.,
RT "The secretory pathway mediates localization of the cell polarity
regulator Alp3p/Bud6p."
RL Mol. Biol. Cell 11:647-661 (2000).
CC -!- FUNCTION: Involved in the organization and/or function of the
actin cytoskeleton.
CC -!- SIMILARITY: TO YEAST BUD6.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z97208; CAB10112.1; --
DR EMBL; AL109770; CAB52420.1; --
DR GeneDB Spombe; SPAC15A10.16; --
DR GO; GO:0005884; C:actin filament; ISS.
DR GO; GO:0016043; P:cell organization and biogenesis; ISS.
DR InterPro; IPR005613; AIP3.
DR Pfam; PF03915; AIP3; 1.
KW Coiled coil, Cytoskeleton.
FT DOMAIN 1009 1096 COILED COIL (POTENTIAL).
SQ SEQUENCE 1385 AA; 154325 MW; ABB3D40CC4FF7537 CRC64;
Query Match 4.3%; Score 163; DB 1; Length 1385;
Best Local Similarity 19.1%; Pred. No. 0.96;
Matches 146; Conservative 129; Mismatches 292; Indels 198; Gaps 33;
QY 26 DLNFQAPMVVTSSTGDL-----SIPSELENIPISENOYFQSAIWS-----GFIKVKKS 74
DB 326 DRKFSPLKTRTPSLTKSLDPGTPSLKSPSLRKSFS--SFVQKDVYSRNSLRISQANRS 383
QY 75 DEVTATSDADNHVTMVVDQEVINKAS-----NSN 104
DB 384 NVFPGATD---VTRSVSHRILSSSTINDGEVAPLPORSTISSPNSPLSATVLPST 439
QY 105 KIRLEKGRIVQIKIYQREN-----PTEKGLDFKL-----YWTDSQNKK 143
DB 440 PILLPRGRSSTLSVKKQFNADGSLTNSPNSIRETEEYVAAKPKLEADIADVETDTSQR 499
QY 144 EVI-----SSDNLQLPELKQKSS-----NSRKKESTAGTVPDRDNGI-----PDSL 187
DB 500 ELLERIQIQAESSEDTSEISLQGLKSLPQVSTQOEIOPSSVPEAASNEIAKEPAVT 559
QY 188 EYEGYTVDVKNKRTFLPSWISNIHEKKGTIKYKSPEKMASTADSDYSPFEKVTGTRDKNV 247
DB 560 AIESIT-----ERKEAPVISSEKIESGTS-STSDTKGGLANFENDSLSELERLIQKN- 613
QY 248 SPAPRPLVAAPYIVHVDNIILSKNEDQSTQNTDSETRTISKNTSRTTSEVHGA 307
DB 614 NAEQDEPSYFKHYKYSSEE-----SGSEDFKSEKDTKGYVISNDSDQVEDESDKSTP 669

QY 308 EVHASFDDIGGSVAGSFNSNSSTVAID-----HSLSLAGERTWAETMGLNTADTARLNAN 363
DB 670 NTGA-----SAKLNDPSSSTITVSDVYKPKASPVEITEPPSSALVSATSPTTNPV 720
QY 364 I--RYVNTGTAPYVNVLPSTSLVLGKNOTLATIDADENQLSQI--LAPNNYYSKNLAPI 419
DB 721 IVPFAVHLSA--FSTAPVSTIV-----SNISPLPTVAPPNVVSGSPSETPI 764
QY 420 ALNAQKFSSTPTMTNYNQFLELEKTKQLRLDQVYGNATYNFENGVRVVDTSNWS 479
DB 765 S-----KPEKVPVVSQTEKAL-----PKPLGVDTIEKYV--FLRYNNQTRKYKVESPLSNAN 813
QY 480 VLPQIOQTARIIPNGK--DLNLVERIAAVNSDPLEITKPDMT-LKEALKIATFGENP 536
DB 914 ELGELFSNVYKISFSGDSYELNIED-----PDTKISYLLDLSLKYKSLVSPFKEQ 866
QY 537 NGNLQ--YQKDIITEFDNFDOOTSONIKNQLAELNATNIYT-VLDKIKLNAKNMLIR 592
DB 867 DANKKREDPHSGEVSATQHSQAQNTLDDHVNTTTHSPSSAPTEILERLKA----- 917
QY 593 DKRFHYDRNNIAGDESVVYKEAHRVINSSTGLLNLNDKDKIRKILSGVIV-EIEDTEG 651
DB 918 -----IEQNIST-----NHTNDSAAALKSSEDSHKLANNFSPDSDIDHKF 956
QY 652 LKEVINDRYDMLNITSSLRQDGKTFI-----DFKK-----YNDK 684
DB 957 YQVQVNMQ-----LELASLKQISAFAFTPIPLKIDPKKEINAFNEK 998
RESULT 11
BXD CLOBO
ID - BXD CLOBO STANDARD; PRT; 1276 AA.
AC P19321;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)
DE (Bontoxilysin D).
GN BOTD.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BVD/-3;
RX MEDLINE=91016853; PubMed=2216736;
RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
RA Kozaki S., Krieselstein K., Henschen A., Gill D.M., Niemann H.;
RT "Nucleotide sequence of the gene encoding Clostridium botulinum
neurotoxin type D."
RL Nucleic Acids Res. 18:5556-5556 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CB16;
RX MEDLINE=93042276; PubMed=1420572;
RA Sunagawa H., Ohyama T., Watanabe T., Inoue K.;
RT "The complete amino acid sequence of the Clostridium botulinum type D
neurotoxin, deduced by nucleotide sequence analysis of the encoding
phase d-16 phi genome."
RL J. Vet. Med. Sci. 54:905-913 (1992).
RN [3]
RP PARTIAL SEQUENCE.
RC STRAIN=D-SA, and D-1873;
RX MEDLINE=89339741; PubMed=2668193;
RA Morishita K., Syuto B., Kubo S., Oguma K.;
RT "Molecular diversity of neurotoxins from Clostridium botulinum type D
strains."
RL Infect. Immun. 57:2886-2891 (1989).
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94230352; PubMed=8175689;
RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,

```
DR Subtilist; BG10797; wapa.
DR InterPro; IPR003305; CBM_Cenc.
DR InterPro; IPR006530; YD_Cenc.
DR Pfam; PF02018; CBM_4_9; 1.
DR Pfam; PF05593; Rhs_repeat; 14.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 17.
KW Cell wall; Repeat; Signal; Complete proteome.
FT SIGNAL 1 28
FT CHAIN 29 2334 WALL-ASSOCIATED PROTEIN.
FT DOMAIN 504 869 3 X 101 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 504 605 1-1.
FT REPEAT 636 736 1-2.
FT REPEAT 769 869 1-3.
FT DOMAIN 1021 2139
FT REPEAT 1021 1040
FT REPEAT 1042 1061
FT REPEAT 1063 1082
FT REPEAT 1083 1102
FT REPEAT 1109 1128
FT REPEAT 1129 1148
FT REPEAT 1150 1169
FT REPEAT 1174 1193
FT REPEAT 1199 1218
FT REPEAT 1219 1238
FT REPEAT 1646 1665
FT REPEAT 1667 1686
FT REPEAT 1690 1709
FT REPEAT 1711 1730
FT REPEAT 1732 1751
FT REPEAT 1753 1772
FT REPEAT 1795 1814
FT REPEAT 1820 1839
FT REPEAT 1840 1859
FT REPEAT 1861 1880
FT REPEAT 1887 1906
FT REPEAT 1908 1927
FT REPEAT 1929 1948
FT REPEAT 1969 1982
FT REPEAT 1983 2002
FT REPEAT 2008 2027
FT REPEAT 2028 2047
FT REPEAT 2051 2070
FT REPEAT 2071 2090
FT REPEAT 2093 2112
FT REPEAT 2120 2139
SQ SEQUENCE 2334 AA; 258329 MW; B75138CCD278BA3 CRC64;

Query Match
Best local Similarity 18.2%; Pred. No. 1.6;
Matches 173; Conservative 126; Mismatches 336; Indels 317; Gaps 42;

QY 1 EVKQENLLNESSSGILGYFSDLNFOA-----PMVVTSTTGDLSIPSSSELENIP 54
Db 269 EVERSDDKSVKLEKNEEGYLLHLTADENWLKDPERVYPSIDPSTLSVSSDTFVMSAYP 328
QY 55 SENYFOQAIWSGFIK--VKKSDEYTPATSADNHVTM-----
Db 329 TTNYSASSQWDAHLKAYVLKGYD-KTGTNYAFMKFNKLPQIQTMTVTKATLKTYVA 387
QY 90 -----WVDDQEVINK-----ASNSKIRLEKGR--LYQIKIQYOR 122
Db 388 HSYGTGTRATGLWLD---TVNSYDNKATWNTKPAKNIQKADVHGQWASVDVTAVK 444
QY 123 ENPTKGLDFKLYWTDSONK---KEVSSDNLQLPKOKSSNRKRSAGTVPY--- 176
Db 445 WNSGGANVGFKLH-TNGNGKEYWKLIS-----SANSANKPYIEVTYTIPTKGN 491
QY 177 -----DRNDGIPD-----SLB-VGQYVDVQNKETFLSPWISNIEHKKGI-TKYKSSP 223
Db 492 TPTTKAYHNGSTGYFDLSWKKVEGAKYKWIYNGKEYQAIAGNV-----TSMSTKG 545
QY 224 EK-WSTASDPDYSDEKVTGRID-KNVSPPEARHPLVAAPYIVHVDMENIILSKNE----- 275
```

```
Db 546 KKIWPTSAETASKRYKL--HLGKQGAELALDP-----SPVYKNSGGSVATSKNYWIGVGA 599
QY 276 --DOSTQNTDSETRTISKNTSTSRHTSEVHGNAEVAHSFEDIGGSVAG-----FS 325
Db 600 IFDQEGEAMSAKPAKPIVNVGKAQAFSAKYNNGNA-TGYFDLSWKAWSGATGYKQOVEN 558
QY 326 NSNSSTVAIDH--SLSLAGERTWAETMGLNTADTA-----RLNANIRYVNTGPA 372
Db 659 GKGFTLDLGNQTSWTTGKKIWIPTSAEIKAGKAYALHLKDGSGAELPINFPGTYKNAGGD 718
QY 373 PIYNVLPITSLVLGKQNTLATIDADENQLSOLLAPNN-----YYPK-----NLA--- 417
Db 719 GAKENYSFKIIAYNKDGEAIAAPATPALPDIAPKPNVTGYLYTNTKSSQGTGVNLINEK 778
QY 418 -----PIALNAQKKFSPTPI-----TMNYNOFLELETK-----QLRL 450
Db 779 VQNAKGYKVIYNGKEYQSFVCGDADHWTTQNKNIWPTSEIKAGSVKYLHTDGKGELAL 838
QY 451 DTDOVTGN-----IATYN--FENGVRVVDVTSNWSSEV-L 481
Db 839 DPSPVYNNANGYKGNYSFTLVAYDANGETPTAPFNPFFHEGAFFLGTBEYWSIIDI 898
QY 482 P--QIQTETARIIFNGKDLNL-----VERRIAAVNPSPDPL-----ETT----- 517
Db 899 PSGQLNGATGVIVNBEEDSIDGPGCLGSRYSNLSDDSLHFGQGWYADAFTSVISTD 958
QY 518 -----KPDMTLKEALKIARG-FNEPNGNLQYQKGIITFDNFDOQTSONIKNQLAEL- 569
Db 959 GGAMYIDEDATTHRTKKADGTYPPTGVYLETETADQFILTKDQTNAYFNKGGKQLQ 1018
QY 570 -----NATNIYTVLDKIKLNAMNLTIRDKRPHYDRN-----NIAVGADSV 611
Db 1019 KVDVGHNAIVYTYNDKQNLTAITDASGRKLITTYDENGHVTSITGPKNKKVTVSYENDL 1078
QY 612 VKE-----AHR--EVINSSTGLLL 629
Db 1079 LKVTDTDGTVTSYDYDSEGRLVKQVSAKSTKAPVFTFYQYSGHRLKAINAKKETYVY 1138
QY 630 NIDKDI-----EKILSGY-----IVEIETDEGLKEVINDRYDMLNI 665
Db 1139 SYDAKDKTLLMTQPNGRKQVGYNEAGNPQIOWIDDAEGLKITTNTKYEKNV 1190

RESULT 10
FAT1_SCHPO
ID FAT1_SCHPO STANDARD; PRT; 1385 AA.
AC O13735; Q9UT0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Actin interacting protein 3 homolog.
GN FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Feat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
```



```

Db 777 YDSYKFEIRKONLGYMGHIVLIAEVVFKLYKVDYISRYSSNLOTETWQYSEVL 836
QY 700 AVTK-----ENTINPSENGDTSN 719
Db 837 NETRMYSKILGGSYIDGNGNIIPQLPDNTTTLTP--NGDASN 880

RESULT 8
SMC3_YEAST
ID SMC3_YEAST STANDARD; PRT; 1230 AA.
AC P47037;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Structural maintenance of chromosome 3 (DA-box protein SMC3).
GN SMC3 OR YJL074C OR J1049.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX NCBI_TaxID=4932;
RN [1]_Yeast;
RP SEQUENCE FROM N.A.
RC STRAIN=W303;
RX MEDLINE=97474309; PubMed=9335333;
RA Michaelis C., Ciolek R., Nasmyth K.;
RT "Cohesins: chromosomal proteins that prevent premature separation of
RL sister chromatids.";
RN Cell 91:35-45(1997).
RP SEQUENCE FROM N.A.
RX MEDLINE=96208490; PubMed=8641289;
RA Galibert F., Alexandraki D., Baur A., Boies E., Chalwatzis N.,
RA Chuat J.-C., Coster F., Cziepluch C., De Haan M., Domdey H.,
RA Durand P., Entian K.-D., Gatius M., Goffeau A., Grivell L.A.,
RA Hennemann A., Herbert C.J., Heumann K., Hilger F., Hollenberg C.P.,
RA Huang M.-E., Jacq C., Jauniaux J.-C., Katsoulou C., Kirchhuth L.,
RA Kleine K., Kordes E., Koetter P., Liebl S.J., Louis E.J., Manus V.,
RA Mewes H.-W., Miosga T., Oberwarter B., Perea J., Pohl T.M.,
RA Portetelle D., Pujoil A., Purnelle B., Ramezani Rad M., Rasmussen S.W.,
RA Rose M., Roesau R., Schaaff-Gerstenschlaeger I., Smits P.H.M.,
RA Scarcez T., Soriano N., Tovan D., Tzermia M., Van Broekhoven A.,
RA Vandenbol M., Wedler D.H., Von Wettstein D., Wambutt R., Zagulski M.,
RA Zollner A., Karpfinger-Hartl L.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RL X.1";
RN EMBO J. 15:2031-2049(1996).
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1678;
RA Sor F.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC1, IRR1 AND MCD1, AND
RX INTERACTION OF THE COHESIN COMPLEX WITH SCC2.
RL MEDLINE=99145468; PubMed=990856;
RA Toth A., Ciolek R., Uhlmann F., Galova M., Schleiffer A., Nasmyth K.;
RT "Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to
RL establish cohesion between sister chromatids during DNA replication.";
RN Genes Dev. 13:320-333(1999).
RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3, MCD1 AND IRR1, AND
RX STRUCTURE.
RL MEDLINE=21980168; PubMed=11983169;
RA Haering C.H., Loeve J., Hochwagen A., Nasmyth K.;
RT "Molecular architecture of SMC proteins and the yeast cohesin
RL complex.";
RN Mol. Cell 9:773-788(2002).
CC -1- FUNCTION: Involved in chromosome cohesion during cell cycle and in
CC DNA repair. Central component of cohesin complex. The cohesin
CC complex is required for the cohesion of sister chromatids after
CC DNA replication. The cohesin complex apparently forms a large
CC proteinaceous ring within which sister chromatids can be trapped.

```

```

CC At anaphase, the complex is cleaved and dissociates from
CC chromatin, allowing sister chromatids to segregate.
CC SUBUNIT: Cohesin complexes are composed of the SMC1 and SMC3
CC heterodimer attached via their hinge domain. MCD1/SCC1 which link
CC them, and IRR1/SCC3, which interacts with MCD1. The cohesin
CC complex also interacts with SCC2, which is required for its
CC association with chromosomes.
CC -1- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatin.
CC Before prophase it is scattered along chromosome arms. At
CC anaphase, the MCD1 subunit of the cohesin complex is cleaved,
CC leading to the dissociation of the complex from chromosomes,
CC allowing chromosome separation.
CC -1- DOMAIN: The flexible hinge domain, which separates the large
CC intramolecular coiled coil regions, allows the heterotypic
CC interaction with the corresponding domain of SMC1, forming a V-
CC shaped heterodimer. The two heads of the heterodimer are then
CC connected by different ends of the cleavable MCD1 protein, forming
CC a ring structure (By similarity).
CC -1- SIMILARITY: Belongs to the SMC family. SMC3 subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; Y14278; CAA74655.1; -
CC EMBL; Z49349; CAA89366.1; -
CC EMBL; X88851; CAA61313.1; -
CC PIR; S56850; S56850.
CC Germline; 141688; -
CC SGD; S0003610; SMC3.
CC GO; GO:0007151; P:sporulation (sensu Saccharomycetes); IMP.
CC GO; GO:0007130; P:synaptonemal complex formation; IMP.
CC InterPro; IPR003439; ABC transporter.
CC InterPro; IPR003405; SMC_C.
CC InterPro; IPR003395; SMC_N.
CC Pfam; PF02483; SMC_C; 1.
CC Pfam; PF02483; SMC_N; 1.
CC Mitosis; Cell cycle; Chromosome partition; ATP-binding; Coiled coil;
CC Nuclear protein.
CC NP_BIND 32 39 ATP (POTENTIAL).
CC DOMAIN 172 482 COILED COIL (POTENTIAL).
CC DOMAIN 483 684 FLEXIBLE HINGE.
CC DOMAIN 685 1041 COILED COIL (POTENTIAL).
CC DOMAIN 1126 1161 ALA/ASP-RICH (DA-BOX).
CC SEQUENCE 1230 AA; 141336 MW; B152D88F7780341F CRC64;
CC Query Match 4.5%; Score 170.5; DB 1; Length 1230;
CC Best Local Similarity 19.8%; Pred. No. 0.34;
CC Matches 182; Conservative 136; Mismatches 355; Indels 247; Gaps 38;
QY 1 EVKQENLLNESBSSQGLLOYFSDLNFOAPMVV--SSTGDLSTPSELENIPSENQ 58
Db 209 EMEQERLEKYNLEERNKIYQFTLYDRLEINQMERLDDGYN-----NTVSYSEQ 262
QY 59 YFQ-----SAIWSGFIKVKSDSEYTPATSDNHNVTWVDDQEVINK--- 99
Db 263 YIQLDKREDMDQVSKLSSI-EASLKIKNATDQQAKLRSEISQKLNVNKKIDVQ 321
QY 100---ASNKNIRLEKGRVYQIK-IQYQRENTPEKGLDFKLVWTSQNKKEVSSDNLQPE 155
Db 322 QQIESNEEQRLNDSATLKEIKSIIEQRKQKLSILPRYQELT---KEEAMY--KLQIAS 375
QY 156 LKQSSNSRKRKSTSGAGTVPDRNDGIPDSLEVEGYTVDKNKRTEFLSPWISNIHE-KK 214
Db 376 LQQQRDLILKGYARFKSKDRDTHSEIE-----ELKSS-----IQNLNELES 422
QY 215 GLTKYKSPKSTASDPYSDFEKVTGRIDKNVSPPEARHPLVAAPIVHVDMEIILSKN 274
Db 423 QLQMDRTSLRKQYSAID-----EEIEELIDISINGPTKGQL-----EDFDSLEHLKQK 471

```

```

Db 810 STDKSTEDNTD---EKYFSAINTN---VTGDSCEIETASNVEE---NLRYCEKD 858
Qy 547 ITEFDENP-DQOTSON---IKNOLA-----ELNATNIYTVLDKIKLNKXNMLIRDKR 595
Db 859 MNEAESGSGDECKQNDGSKTKQISFTSDPNFOESNDTEFSSTK----- 905
Qy 596 PHYDRNNIAGADESVVKEAHR-EVIN-----SSTEGLLINIDKD 634
Db 906 ---YKVRNSDLEDESLSKELTKAEVVVDKLBDEESEDSEYQDYADPEPGNDEGSNENIVKG 963
Qy 635 IKKILSGYIVETDEGLKEVINDRYDMLNISLQDQKTFIDFKYNDKPLVLSNPNY 694
Db 964 TKK-----DTLGIYEPENKVN-----KVHEETLFEANVSS 995
Qy 695 KVNIVAVTKENTIIN 709
Db 996 SVNVQNKDHTDVIN 1010

```

RESULT 7

```

S155 YEAST
AC P436T2;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S1T4-associated protein SAPI55.
GN SAPI55 OR YFR040W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RN MEDLINE=96220458; PubMed=8649382;
RA Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA Arndt K.T.;
RT "The SAP, a new family of proteins, associate and function positively
RT with the S1T4 phosphatase."
RL Mol. Cell. Biol. 16:2744-2755 (1996).
[2]
SEQUENCE OF 98-1000 FROM N.A.
RN STRAIN=S288C / AB972;
RC MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RL Nat. Genet. 10:261-268 (1995).
[3]
SEQUENCE OF 98-1000 FROM N.A.
RN STRAIN=S288C / AB972;
RC MEDLINE=95287652; PubMed=8686379;
RA Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,
RA Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of
RT chromosome VI from Saccharomyces cerevisiae."
RL Yeast 12:149-167 (1996).

```

```

CC -1- FUNCTION: Associates with the S1T4 phosphatase in a cell cycle
CC dependent manner. May be directly or indirectly involved in
CC S1T4-dependent functions in budding and in normal G1 cyclin
CC expression.
CC -1- PPM: Hyperphosphorylated in the absence of S1T4.
CC -1- SIMILARITY: Belongs to the SAPS family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; U50560; AAC49303.1; -.
DR EMBL; D50617; BAA09279.1; -.
DR GenBank; 140194; -.
DR SGD; S0001936; SAPI55.
DR InterPro; IPR007587; SAPS.
DR Pfam; PF04499; SAPS; 1.
DR Phosphorylation; Cell cycle.
FT CONFLICT 663 689
FT CONFLICT 674 689
FT CONFLICT 814 815
FT CONFLICT 818 822
FT CONFLICT 822 822
SQ SEQUENCE 1000 AA; 114880 MW; AD4FBF0BC7C588D5 CRC64;

```

Query Match 4.5%; Score 171; DB 1; Length 1000;

Best Local Similarity 19.5%; Pred. No. 0.25; Matches 173; Conservative 130; Mismatches 321; Indels 262; Gaps 42;

```

Qy 1 EVQENRLLNESSSSQGLLGYYFS--DL-----NFQAPMVVTSTTG--LSIPS 47
Db 90 EKEEERSMNSSESSTSFSSGSTSKTDLDEEDISNATAPMVVTYTNKLNDSFIERLVLVET 149
Qy 48 SELENTPEENQYFQSALWSGFIKVKKSDEYTFATSNADNHVWVDD-QEVINKASNSKI 106
Db 150 ELNELSRQNKTLIDPICFGFFDKTKNKV-----NNWEYLVDQLMCEISKIKTATTV 203
Qy 107 RLEKGRLYQIKYQRENPTKGLDFKLYWTDSONKEVSSDNQLQPELKQKSSNRKK 166
Db 204 DLNN-----LIDYQEQQLD-----DSSQEDVYVESDEQEKEEDDNNNNKK 247
Qy 167 R-----STSAGPTVPEDNDGIPDSLEVBGYTVDVKNKRTFLSP-----WISNIHEKKG 215
Db 248 RRRKGSFNGDDNNNDNDDDDANDEDESAYL-----TKATIISEIFSLDIWLIS-----ES 299
Qy 216 LTKYKSSPEK-WSTASDPYSDFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKN 274
Db 300 LVKNQSYLNKWSIINQF-----NFNSE-NSPLV---DIFLKNQNLTLTR- 341
Qy 275 EDQ-----STQNTDSET-----RTISKNT 293
Db 342 QDQYLNFIPTERSFVDDMLKHVDISLLMDFLKIISTDKIESPTGIIELVYDQNLISKCL 401
Qy 294 STSRHTSEVHGNAEVAHFDDIGGSVAGFSNSNSTVAI-DHSL--LAGERTWAETM 350
Db 402 SFLNNKSP---ADIQACVGDFLKALIAISANAPLDDISIGPNSLTROLASPESIAKL 457
Qy 351 GLNTADTARLNANIRYVNTGTAPIYVLPPTSLVLGKNQTLATIDADENQSLQILAPNNY 410
Db 458 DI-----MINQGAALNT-----TVSIVI---ELIRKNSDYDQVNLITTTIKT 498
Qy 411 YPSKNLAPIALNAQKFSSTPTMNYNQFLEKTKOLRLDTQVYGNATYNPENGRVR 470
Db 499 HPPSNRDPYLYGLLRKFSNHLSDFFQIILDIENDANIPLENQLHEKFKPLGPERKV- 557
Qy 471 VDTGSNNSEVLPOQTETARIIFNGKDLNIVERRIAAVNP--SDPLETTPKDMTL--KEA 526
Db 558 -----VELIAELLHCNNMGLMNSKRAERIAARRDKVRSQLSHLLQDALNDLSIEKEQ 610
Qy 527 LKIAFG-----FNEPNGNLQYQKDI-----TEPDFNP--QQTSQNIKNOLAE 568
Db 611 LKTKHSPTRTDHLKNNNGKINDNNDNDDSDYGDDEIDSEFPIYNNKQNLKRLTDP 670
Qy 569 LNATNIYTVLDKIKLNARMLIRDKRFHYDRNNIAVGAD-----ESVVKAEHREVIN-- 621
Db 671 TVGT-----YKSNYMIKD-----CFQNNELFLTHPWNFWHNVIFDIIOQIFNGR 716
Qy 622 ---SSTEGLLINID--KDRIKILSGYIVET---DTEGLKEVINDRYDMLNLS--LR- 669
Db 717 MDFSNGFLVLSLNLKSSYQFMVDIVISDEKGTDSVRSFVRDPNDFPKITTFILRG 776
Qy 670 -QDGKTFIDFKKNY-----DKLFLYISN-----PNYKNVNY 699

```

```

103 SNKIRLEKGLYQIKYQYRENTPEKGLDPLKFWTDSQNKKEVISDNQLQLBELKQKGSN 162
168 TSKYKLDKEBAQKLLIR-----TNDALD-KLOGAIKELERQVNSLDK-QASAKIYLEK 219
163 SRKRSSTASPTVDRD-----NDGIPDSLEVEGTVDVK-NKRTFLSPWISNIHEKG 215
220 SKALESVEGLIYVNDLNFNEKLNNTLSLEVEQQRNDLEINIQYESSISQTVH----- 275
216 LTKYSPEKWSPTASDPYDFEKTGRID--KNVSP-----EAPHLVAAYPIVHV 264
276 ---FKTEVE-----SSIQBITSKLDNLKNALSEINLQEARIBERRKLIISGEIV-V 322
265 DMENII--LSKN-EDQSTQNTDSETRISKNSTSTHT-----300
323 DQTKIEIKKQVESLKIQINAKQREIELDQQLTRLNKANKSLKLOENDINKGCVLLE 382
301 --SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAE---TWGLNTPA 355
383 KGSAAANINILKQPFENKFSKGIKTIDKNSFLFDGVIGLASELFPKVESEFSLAIVT 442
356 DTAELN-----ANRYVNTGTAPINVLPTSL-----VLG 386
443 LGAALNOIVMKTSDVLQADFLKKNLSGKATFIPTSIKEREVRDHLVLKQKGFUG 502
387 KNQTLATIDADENO-----LSQILAPNNYYPSPKXLAPI-----ALN 422
503 VAKELIEFDQFNKLFGLLGNILVVDVNDVNAKRIKILDKHTIVSLEGLDFRPGGTIT 562
423 AQKFSPTPTMNTNQFLEKTKQLRLDQD-----VYGNATYVNFENGRVVD 472
563 GGSKLERTSI--LNYDIKIK-EHTNLTAFEDQIHLKIKQOTIYNEIETVNSTIQOVKIE 620
473 TGS-----NWSVLPOIQE--TTARIIFNGKDLNLVERRIAAV-----508
621 ANSINKLINLEBNLKNLNASFIFKEQEDQESLNSFDSEKLN-IEKQISTLTIELN 679
509 NPSDPL-----ETTKPDWTLK-----EALKIAGFNEPFGNLOVQOK 545
680 SKDRLTLNISEQKGETTKQELDAKRLKLTQHSDSITEQNAKFLVEQNKRLSEHYK 739
546 DITE-----FDNFDOOTS--ONIKNQLAELNATNIVTVLDKIKLNAXMNLIRDKRF 596
740 LTLNASEQYSLDLDIEQARHFVDSLKELGELGNVLEAITFEFVYVQR-----YQEKQ 795
597 HYDRNNIAGDESVVYEAHREVINSSTEGLLNIDKD-----IRKLSGYVIEIEDTEG 651
796 YIEELTTAKSIEBAISDLXIIINKTTE--IVNLVNEFNFMVQKMGFGGKAEIHFT-- 851
652 LKEVINDRYDMLNIS---SLRQDGKTFIDFKKYND-----KLPLYI- 689
852 -----DKNDILNGSVEISAQPPGKTINKLRLFGSGEKALIAISLLFALKARPILCIL 905
690 -----SNPNYKNVNYAVTKENT 706
906 DEVEAALDESIVRYVEFLKULKENT 931

RESULT 6
YM67 YEAST STANDARD; PRT; 1658 AA.
AC Q03661; Q04988;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 187.1 kDa protein in GUA1-BRG8 intergenic region.
GN YMR219W OR YMR261.13 OR YMR59.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STAIN=S288C / AB972;
RA MEDLINE=97333268; PubMed=9169872;
RA Bowman S., Churcher C.M., Baddock K., Brown D., Chillingworth T.,
RA Connor R., Dedman R., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII."
RL Nature 387:90-93(1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z49809; CAAB9934.1; -.
DR EMBL; Z49939; CAAG0190.1; -.
DR PIR; S55101; S55101.
DR GerMOnline; 142894; -.
DR SGD; S0004832; ESC1.
DR GO; GO:0005634; C:nucleus; IDA
DR GO; GO:00056348; P:chromatin silencing at telomere; IMP.
DR Hypothetical protein.
KW SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;
SQ
Query Match 4.6%; Score 173; DB 1; Length 1658;
Best Local Similarity 18.4%; Pred. No. 0.39;
Matches 157; Conservative 140; Mismatches 278; Indels 280; Gaps 40;
QY 1 EVKQENRLNSES---SSQGLLYYFSDLNLFQAPVWVTSSTGDLSPSELENISEN 57
Db 290 EMELEDDIDVESDAEKDESQAGTEHS-VDFSYMQPRTDNTKIPVIEKSEDEHKVHQ 348
QY 58 QYFQSAIWS-GFIKVKKSD-----YTFATSAQN---HVTWVDDQDEVINKASNSNK 105
Db 349 RISEDCATDFGSVIVSDDESEDESAESYSAENAVYHNEHELDKLELIEDSDS 408
QY 106 ILEKGRLYQIKYQYRENTPEKGLDPLK-----YTTDSQNKKEV 145
Db 409 -----ESQSAESEQGSEDDPFYKMKVEKSTSEETENTSESRDQGFADAYTKKV 459
QY 146 ISSDNQLQLPEL-----KOKSSNRKGRSTASGPTVPDRDNGIPDSLEVE 190
Db 460 EQQENDERPEKDDIIRSLDKNFHGNKNKSEYSENVLENETDPAIVERNQ-IND---VE 515
QY 191 GYTVDVQNKRTFL---SPWISNIHE--KKGLTKYKS-----PEKWTASD----- 231
Db 516 GYDVTGKSVESDLHEHSP--DNLYDLAARAMLQFQSRNSNCPQKEQVSESYLHSGNS 573
QY 232 -----PYSDFEKTGTIDKQVSPFARHPLVAAYPIVHVDMENIILSKNEDQ 277
Db 574 NLSGRSLDESEEQIPLKDF---TGNNNNLKT-D-RGDLSSS---VEIEVEKV-----SEKK 622
QY 278 STQNTDSETRITISKNTSTSRHTSEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHS 337
Db 623 LDGSTKELVPLSTDTTN-----NSSLGNEDSIYYSLDD- 657
QY 338 LSLAGERTWAEVTMGLNTADTALNANIRYVNTGTAPINVLPTTSVLKCKNCTLATIDAD 397
Db 658 -----ADAISENLTD-----VPLMEIKTTPKYEVWISESV-----YSSTSYE 694
QY 398 ENOLS-----QILAP-----NNYYPSPKN-----LAPIALNAQKFSSTPTIM 434
Db 695 DNTVAMPQQVEYTSFPMNDPFNSLNDYKGDHLKSTLALAPAFKKADEFVEAGVTK 754
QY 435 -----NYNQFLEKTKQLRLDQVYGNIAIYFNENGRVVDTSNWNSEVLPTQIE 486
Db 755 SCLTSTSGHTNIFHTSKETKQVS-DLDESTENVTTFENENTG----DENKQSKNTPGVAN 809
QY 487 TTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDWTLKEALKIAGFNEPFGNLOVQOKD 546
```

DR EMBL: X02919, CAA26676.1, --
 DR PIR: A24594, A24594.
 DR PIR: S05603, S05603.
 DR PDB: 1CEJ, 28-MAY-99.
 DR InterPro: IPR006209, EGF_like.
 DR Pfam: PF00008, EGF, 1.
 DR Malaria: Merzoite; Polyprotein; Repeat; Signal; Glycoprotein;
 KW Transmembrane; GPI-anchor; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 1639
 FT MERZOITE SURFACE PROTEIN 1.
 FT FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 764 764 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 768 768 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 783 783 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 844 844 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 920 920 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;
 Query Match 5.0%; Score 190.5; DB 1; Length 1639;
 Best Local Similarity 20.6%; Pred. No. 0.051;
 Matches 171; Conservative 138; Mismatches 255; Indels 265; Gaps 47;
 QY 5 ENRL-LNESSESSQGLGYFSDLNFPQAPWVTSSTTGLSPSELENIPENQVPSQA 63
 DB 927 ENLISLGNKYNQELIGQSSS-NF-----YKILKDSFTFYNE 965
 QY 64 IWSGFIVKKSDEYTFATSDNHNVTWVDDQEVINKASNS-----NKILEKGLRY- 114
 DB 966 SFINFVKSADD-----INSLNDESKRKLLEDINKLKTQLSPDLNKKYKLERLFD 1020
 QY 115 -----QIKIQORENPTKGLDFKLYWTDQSNKKEVIVSDNLQDELKQKSNRKRKS 168
 DB 1021 KKTATGKYMQIKKLLKQESKL--NSLANPRHL--QNFVFPNKKAEAEATEN 1076
 QY 169 TSAGPTVPDRDNGIPDSLEVEGYTVDKNKRFLSPWISNIEKKGLTKY---KSSP-- 223
 DB 1077 T-----LENTKILLKHY-----KGLVYKNGSSPLK 1103
 QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPPEAR-----HPLVAAYPIHVVD 265
 DB 1104 TLSEESIQTEDNVASLENFKVLKLEGLKDLNLEKKLSYLSGLHLHIA-----E 1156
 QY 266 MENILSKNEDQSTQNTDSETRITKNTSTRTHTSEVHGNAEVHASFDDIGSVSAGFS 325
 DB 1157 LKEVIKKN---YTGNSPSENNT-----DVNALESYKKFLPEGTIDVATVVS 1200
 QY 326 NSNSSTVAIDHSLSLAGERTWAETMGLTADTARLNANIRYVNTGAPTINYVLPPT--- 381
 DB 1201 ESGSDTLEQSQPKKPASTHVGAES---NTITTSQ-NVDDVEDVDVIIPIFGSEEDYDDL 1256
 QY 382 -SLVLGKNTLATIDADENQLSOLAPNNYPSKNLAPIA---LNQKXFSPTPTMYN 437
 DB 1257 GQVVTGEAVTPSVID---NLSKI---ENEYEVLYLKLPLAGVYRSLSKGLNNVMTFVN 1310
 QY 438 -----QFLEKTKQLRLDTQV-YGNIAT-----YNFENGRVRVDTGNSWSEVL 481
 DB 1311 VKDILNSRFKRENFKNV-LESDLIPYKOLTSNYYVVPKFKLNKEDKFLSSNYIK 1369
 QY 482 PQIQTETARIIFNGKDLNLVERIAAVNPSDPI-----ETTKPDM-TLKALKIAPGF 533
 DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSKSLDSIKKIYINDKQGE 1410
 QY 534 NEPN-----GNLYQGGKDIET-----FDNFQQTSONIKNOLAELNATNIYT 576
 DB 1411 NEKYLPLNIEITLYKTVDNKIDLFIHLEAKVLYNTYK---SNVEVIKELN---YLKT 1465

QY 577 VLDKIKLNKNN--ILIRDKRFHYDRNN-----IANG-ADSVVKEAHREVINSSTEGLL 628
 DB 1466 IQDKLADFKKNNFVGIALSTDYNNHNLTKFLSTGMVFENLAKTVLSNLLDGNLQGG-M 1524
 QY 629 LNIDKD--IRKILSGYIVETEDTEGLKEVINDRYDMLNLSLQQDQGTFFDFKKYNDKLP 686
 DB 1525 LNIQHQCVRK-----QCPQNSGCFRLHDE-----RECKCLLNKQEGDKC- 1566
 QY 687 LYISNPNYKNVAVTENTIIINPENG-----DTSTNGIKKI 724
 DB 1567 --VENPNPTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1599
 RESULT 5
 ID P115_MYCHR STANDARD; PRT; 979 AA.
 AC P41508;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE P115 protein.
 OS Mycoplasma hyorhinis.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2100;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=9113890; PubMed=1825306;
 RA Notarnicola S.M., McIntosh M.A., Wise K.S.;
 RT "A Mycoplasma hyorhinis protein with sequence similarities to
 RT nucleotide-binding enzymes.";
 RL Gene 97:77-85(1991).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- DOMAIN: Consists of two putative central coiled-coil regions
 CC flanked by putative globular regions at the N- and C-termini.
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYOPLASMA
 CC P115.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M34956; AAA25423.1; -.
 DR PIR: JQ0894; JQ0894.
 DR InterPro: IPR003439; ABC transporter.
 DR InterPro: IPR005289; GTP-binding_dom.
 DR InterPro: IPR003405; SMC_C.
 DR InterPro: IPR003395; SMC_N.
 DR Pfam: PF02483; SMC_C; 1.
 DR Pfam: PF02463; SMC_N; 1.
 DR TIGRFAMs: TIGR00650; MG442; 2.
 DR ATP-binding; Coiled coil.
 KW NP_BIND 32 39 ATP (POTENTIAL).
 FT DOWAIN 169 224 COILED COIL (POTENTIAL).
 FT DOWAIN 231 400 COILED COIL (POTENTIAL).
 FT DOWAIN 569 821 COILED COIL (POTENTIAL).
 FT DOWAIN 884 912 ALA/ASP-RICH (DA-BOX).
 SQ SEQUENCE 979 AA; 110566 MW; 30D51C56B56280F4 CRC64;
 Query Match 4.6%; Score 174.5; DB 1; Length 979;
 Best Local Similarity 19.1%; Pred. No. 0.16;
 Matches 177; Conservative 132; Mismatches 134; Indels 277; Gaps 40;
 QY 13 ESSSQGLGYFSDLNFPQAPMV-----TSSTTGLSPSELENIPSE 56
 DB 51 EQSAQLRGLNMDVIFAGSKTVKPKQKAMVLTFFKNEDAIEETKQFTISRLKRGQGT 110
 QY 57 NOYFQS-----AIMSGFIKKSDEYTFATSDNHNVTWVDDQE--VINKASN 102
 DB 111 NEYFNDQPVRYKDKNLAVESG---ISKSLAISQGTISEIAETPEOKAVIEEAG 167

RT "Polymorphism of the precursor for the major surface antigens of
RT Plasmodium falciparum merozoites: studies at the genetic level."
RL EMBO J. 4:3823-3829(1985).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RA Fan W., Tolle R., Bujard H.
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42
CC kDa and 19 kDa antigens which are the major surface antigens of
CC merozoites. The maturation take place during schizont.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X03371; CAA27070.1; -
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL. SURFACE PROTEIN 1.
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDE3CE0A46322 CRC64;
Query Match 5.0%; Score 190.5; DB 1; Length 1630;
Best Local Similarity 20.6%; Pred. No. 0.05;
Matches 171; Conservative 138; Mismatches 255; Indels 265; Gaps 47;
QY 5 ENRL-LNESSSSSGGLLYFSDINFOAPMVVTSSTGDLSTPSSSELENIPSENQYFQSA 63
DB 918 ENILSLGKNKNIYQELIGQSSSE-NF-----YKILKDSDTFYNE 956
QY 64 IWSGFIVKKSDEYTFATPSADNHTVMVDQEVINKASNS-----NKIRLEKRLY- 114
DB 957 SFTNFVSKADD-----INSINDESKKLEEDINKLKTQLSFDLYNKYKLERLFD 1011
QY 115 -----QIKTYQRENTEKGLDFKLYWTDQSNQKKEVISDNILQPLKQKSNRSKRS 168
DB 1012 KKTGVGKYKQIKKLTLLKEQLESKL--NSLNNPKHVL--QNFSPVFNKKAETAEEN 1067
QY 169 TSAGTPYDRDNDGIPSLSEGVGVVDVKNKRTFLSPWISNIHEKKGLTKY--KSSP-- 223
DB 1068 T-----LENTKILLKHY-----KGLVKYNGESSPLK 1094
QY 224 ---EKWSTADSPYSDFE-----KVTGRIDKNVSPPEAR-----HPLVAAYPIVHVD 265
DB 1095 TLSEESITQEDNVASLENFVKLSKLEGLKDNLNLEKKLSYLSGSLHLHIA-----E 1147
QY 266 MENILSKNEDQSQNTDSTRTISKNTSRTHTSVHGNARVHASEFDIGSVAGFS 325
DB 1148 LKEVINKN-----YTGNSPSENNT-----DVNNALSEYKKFLPEGTDVATVVS 1191

QY 326 NSNSSTVAIDHSLSLAGERTWAEGLNTADTARLNANRYVNTGTAPYVLPPT--- 381
DB 1192 ESGSDTLEOSQPKKPASTHVGABS---NTITTSQ-NVDDEVDDIIVPFIQSEEDYDDL 1247
QY 382 -SLVLGKNOTLATIDADENOLSOILAPNNYPSKILAPIA---LNAQKFFSSTPIITMNTN 437
DB 1248 GQVVTGEAVTPSVID---NILSKI---ENEYEVLYLKPLAGVYRSLSKKOLENNVMTFVN 1301
QY 438 -----QFLELEKTKQLRLDQV- YGNIAI-----YNFENGRVAVDTGSGNWSVL 481
DB 1302 VKDILNSRFKRENPKNV-LESDLIPYKDLTSSNYVVKDPYKPLNKEKDKFLSSYNYIK 1360
QY 482 PQOETIARTIIFNGKDLNLVERRIAAVNPSPDL-----ETTKPDM-TLKALKIAPGF 533
DB 1361 DSID-----TDINFA-----NDVLGYKILSEKYKSLDLSIKKVINDKQGE 1401
QY 534 NEPN-----GNLOYQGDITE-----FDNFQDQTSQNTKNOJLAELNATNIYT 576
DB 1402 NEKYLPPLANNIETLYKTVDNKIDLFVHLEAKVLNITYEK---SNVEVIKELN--YLKT 1456
QY 577 VLDKIKLNKKN--ILIRDKRFHYDNN-----IAGV-ADSVVKEAHREVINSSTEGLL 628
DB 1457 IQDKLADFKNNNFVGIADLSTDYNNHLLTKFLSTGMVFENLAKTVLSNLLDGNLQGG-M 1515
QY 629 LNIDKD--IRKILSGYVIEIETDEGLKVINDRYDMLNLSLRQDGKTFIDFKYNDKLP 686
DB 1516 LNISQHCQVKK-----QCPQNSGCFRLDE-----RECKCLLYKQEGDKC- 1557
QY 687 LYSNPNYKVNVAVTYKENTIIINPSENG-----DTSNNGIKKI 724
DB 1558 --VENPNPTCN-----EN-----NGGCDADAKTEEDSGNSG-KKI 1590
RESULT 4
MSPI_PLAFW STANDARD; PRT; 1639 AA.
ID P04933;
AC 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PUMSA) (P195).
GN MSP-1.
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5948;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86014355; PubMed=2995820;
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
RA Nicholas S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
RA Freeman R.R.;
RT "Primary structure of the precursor to the three major surface
RT antigens of Plasmodium falciparum merozoites."
RL Nature 317:270-273(1985).
RN [2]
RP REVISIONS.
RA Holder A.A.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42
CC kDa and 19 kDa antigens which are the major surface antigens of
CC merozoites. The maturation take place during schizont.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----


```
Query Match 99.2%; Score 3745; DB 1; Length 764;
Best Local Similarity 99.3%; Pred. No. 2.7e-180;
Matches 730; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFDLNFQAPMVVTSSTTGDLSIPSELENISENQYF 60
DB 30 EVKQENRLNSESSESSQGLGYFDLNFQAPMVVTSSTTGDLSIPSELENISENQYF 89
QY 61 QSAIWSGFIKVKSDYEYFATSAADNHVMTWDDQVINKASNSKIRLEKRLVQIKIQY 120
DB 90 QSAIWSGFIKVKSDYEYFATSAADNHVMTWDDQVINKASNSKIRLEKRLVQIKIQY 149
QY 121 QRENTEKGLDFKLYWTSDQNKKEVYISDNLQLPKQKSNRKRSTAGTPVPDRDN 180
DB 150 QRENTEKGLDFKLYWTSDQNKKEVYISDNLQLPKQKSNRKRSTAGTPVPDRDN 209
QY 181 DGIPDSLEVEGTVTVKQKRTFLSWINIEHKKGLTKYKSPKWSASTADPYDFEKT 240
DB 210 DGIPDSLEVEGTVTVKQKRTFLSWINIEHKKGLTKYKSPKWSASTADPYDFEKT 269
QY 241 GRIDKXNVSPEARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
DB 270 GRIDKXNVSPEARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 329
QY 301 SEVHGNAEYHASFDFIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEETMGLNTADTARL 360
DB 330 SEVHGNAEYHASFDFIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEETMGLNTADTARL 389
QY 361 NANIRVNTGTAPIYVNLPTTSLVIGKQTLATIDADENQSLQILAPNNYPSKMLAFIA 420
DB 390 NANIRVNTGTAPIYVNLPTTSLVIGKQTLATIIKAKENQSLQILAPNNYPSKMLAFIA 449
QY 421 LNAQKFFSTPTIMTYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
DB 450 LNAQKFFSTPTIMTYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 509
QY 481 LPOIQETTAIRIIFNGKDLNVERIAAVNPSPLETTKPDMTLKEALKIAGFNEPGNL 540
DB 510 LPOIQETTAIRIIFNGKDLNVERIAAVNPSPLETTKPDMTLKEALKIAGFNEPGNL 569
QY 541 QYQKDIITEFDNFDOOTSONIKNQLAELNATNITVLDKIKNAKMLIRDKFHYDR 600
DB 570 QYQKDIITEFDNFDOOTSONIKNQLAELNATNITVLDKIKNAKMLIRDKFHYDR 629
QY 601 NNIAGADSVVKEAAREVINSSTEGLLNIDKIRKILSGYVIEDTEGLKEVINDRY 660
DB 630 NNIAGADSVVKEAAREVINSSTEGLLNIDKIRKILSGYVIEDTEGLKEVINDRY 689
QY 661 DMLNISLQDQKTFIDFKYNDKPLYSNPYKVVAVTKENTIIINPSENGDTSTNG 720
DB 690 DMLNISLQDQKTFIDFKYNDKPLYSNPYKVVAVTKENTIIINPSENGDTSTNG 749
QY 721 IKKILIFSKGYEIG 735
DB 750 IKKILIFSKGYEIG 764

RESULT 2
YPBI_BACAN STANDARD; PRT; 204 AA.
ID YPBI_BACAN
AC P13422; Q9X377;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein pX01-111.
GN pX01-111.
OS Bacillus anthracis.
OC Plasmid pX01.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.

QY MEDLINE=89172073; PubMed=3148491;
RA Walkes S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppla S.H.,
RA Schmidt J.J.;
RT "Sequence and analysis of the DNA encoding protective antigen of
RT Bacillus anthracis."
RL Gene 69:287-300(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=sterne;
RX MEDLINE=99445483; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
RA Keim P., Koeber T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
RA Martinez Y., Rieke D., Svensson R., Jackson P.J.;
RT "Sequence and organization of pX01, the large Bacillus anthracis
RT plasmid harboring the Anthrax toxin genes."
RL J. Bacteriol. 181:6509-6515(1999).
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22589; AAA22636.1; -
CC EMBL; AF065404; AAD32415.1; -
CC PIR; G59104; G59104.
CC PIR; I39933; I39933.
CC HSP; P13423; JACC.
KW Hypothetical protein; Plasmid; Transmembrane.
FT TRANSMEM 162 182 POTENTIAL.
FT CONFLICT 93 93 T -> A (IN REF. 1).
FT CONFLICT 184 204 KSCNCLIVYEVSQLNSVFI -> NHVIVYLSM (IN
FT REF. 1).
SQ SEQUENCE 204 AA; 23029 MW; E1657B23AE4273FD CRC64;

Query Match 6.2%; Score 235.5; DB 1; Length 204;
Best Local Similarity 34.6%; Pred. No. 1.7e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNTILDRPHYDNNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYVIEI 646
DB 1 MNTILVRDP-YYDNGNIVGVDDSYLKNAYKQILNWSGSLNLDQVQALSGYMLQI 59
QY 647 EDTE-----GLKEVINDRYDMLNISLQDQKTFIDFKYNDKPLYSNP 693
DB 60 KKPSENHLTNSGFTITLAGKDSGVGELYRLVS-----DGTGFLDFNKFENWRSLV-DPG 112
QY 694 YKVVVAVTKEN-TIINPSENGDTSTNGIKKILIFSKGYEII 734
DB 113 DDVVVAVTKEDFNATRDENGNTA-NKLNKTLVLSGKIKEI 153

RESULT 3
MSP1_PLAFK STANDARD; PRT; 1630 AA.
ID MSP1_PLAFK
AC P04932;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMWASA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman M., Bore N., Hyde J.E., Scaife J., Certa U.,
RA Stunnenberg H., Bujard H.;
```

J. Biol. Chem. 266:15493-15497(1991).
[6] CHARACTERIZATION.
RC STRAIN=Sterne;
RX MEDLINE=94327640; PubMed=8051159;
RA Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;
RT "Anthrax protective antigen forms oligomers during intoxication of
mamalian cells.";
RL J. Biol. Chem. 269:20607-20612(1994).
[7] CHARACTERIZATION.
RP MEDLINE=21129592; PubMed=11207581;
RX Bearegard K.E., Collier R.J., Swanson J.A.;
RA "Proteolytic activation of receptor-bound anthrax protective antigen
on macrophages promotes its internalization.";
RL Cell. Microbiol. 2:251-258(2000).
[8] TOXIN REGULATION.
RP STRAIN=Weybridge;
RX MEDLINE=94131936; PubMed=8300513;
RA Koehler T.M., Dai Z., Kaufman-Yarbray M.;
RT "Regulation of the Bacillus anthracis protective antigen gene: CO2 and
a trans-acting element activate transcription from one of two
promoters.";
RL J. Bacteriol. 176:586-595(1994).
[9] MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.
RP STRAIN=Sterne;
RX MEDLINE=95050722; PubMed=7961869;
RA Singh Y., Klimpel K.R., Aroxa N., Sharma M., Leppla S.H.;
RT "The chymotrypsin-sensitive site, PFD315, in anthrax toxin protective
antigen is required for translocation of lethal factor.";
RL J. Biol. Chem. 269:29039-29046(1994).
[10] MUTAGENESIS OF DOMAIN 4 LOOPS.
RP STRAIN=Sterne;
RX MEDLINE=99185012; PubMed=10085028;
RA Varughese M., Teixeira A.V., Liu S., Leppla S.H.;
RT "Identification of a receptor-binding region within domain 4 of the
protective antigen component of anthrax toxin.";
RL Infect. Immun. 67:1860-1865(1999).
[11] MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.
RP STRAIN=Sterne;
RX MEDLINE=21092804; PubMed=11178978;
RA Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;
RT "Trp 346 and Leu 352 residues in protective antigen are required for
the expression of anthrax lethal toxin activity.";
RL Biochem. Biophys. Res. Commun. 281:186-192(2001).
[12] MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.
RP STRAIN=Sterne;
RX MEDLINE=21438996; PubMed=11554763;
RA Ahuja N., Kumar P., Bhatnagar R.;
RT "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are
required for oligomerization of anthrax protective antigen.";
RL Biochem. Biophys. Res. Commun. 287:542-549(2001).
[13] MUTAGENESIS OF PRO-289.
RP STRAIN=Sterne;
RX MEDLINE=21253689; PubMed=11356563;
RA Khanna H., Chopra A.P., Arora N., Chaudhry A., Singh Y.;
RT "Role of residues constituting the 2betaal strand of domain II in the
biological activity of anthrax protective antigen.";
RL FEMS Microbiol. Lett. 199:27-31(2001).
[14] MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.
RX MEDLINE=21125576; PubMed=1122612;
RA Mogridge J., Moutez W., Collier R.J.;
RT "Involvement of domain 3 in oligomerization by the protective antigen
molecy of anthrax toxin.";
RL J. Bacteriol. 183:2111-2116(2001).
[15]

MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.
RX MEDLINE=21269403; PubMed=11113126;
RA Sellman B.R., Nassi S., Collier R.J.;
RT "Point mutations in anthrax protective antigen that block
translocation.";
RL J. Biol. Chem. 276:8371-8376(2001).
[16] MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;
RX ILE-239; TRP-255 AND PHE-265.
RC STRAIN=Sterne;
RX MEDLINE=22112896; PubMed=12117959;
RA Chauhan V., Bhatnagar R.;
RT "Identification of amino acid residues of anthrax protective antigen
involved in binding with lethal factor.";
RL Infect. Immun. 70:4477-4484(2002).
[17] X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RP MEDLINE=97192099; PubMed=9039918;
RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;
RT "Crystal structure of the anthrax toxin protective antigen.";
RL Nature 385:833-838(1997).
[18] REVIEW.
RP MEDLINE=21428689; PubMed=11544370;
RA Mock M., Fouet A.;
RT "Anthrax.";
RL Annu. Rev. Microbiol. 55:647-671(2001).
[19] FUNCTION: One of the three proteins composing the anthrax toxin,
the agent which infects many mammalian species and that may cause
death. PA binds to a receptor (ATR) in sensitive eukaryotic
cells, thereby facilitating the translocation of the enzymatic
toxin components, edema factor and lethal factor, across the
target cell membrane. PA associated with LF causes death when
injected. PA associated with EF produces edema. PA induces
immunity to infection with anthrax.
[20] SUBUNIT: Anthrax toxins are composed of three distinct proteins, a
protective antigen (PA), a lethal factor (LF) and an edema factor
(EF). None of these is toxic by itself. PA+LF forms the lethal
toxin (LeTx); PA+EF forms the edema toxin (EdTx). PA-63 forms
heptamers and this oligomerization is required for LF or EF
binding. Once activated, at low pH, the heptamer undergoes
conformational changes and converts from prepore to pore inserted
in the membrane, forming cation-selective channels.
[21] SUBCELLULAR LOCATION: Secreted.
[22] DOMAIN: The molecule is folded into four functional domains. Each
domain is required for a particular function in the toxicity process.
Domain 1 contains two calcium ions and the proteolytic activation
site. Cleavage of the PA monomer releases the subdomain 1a, which
is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is
part of the remaining 63-kDa fragment (PA63) and contains the
binding sites for LF and EF. Domain 2 is a beta-barrel core
containing a large flexible loop that has been implicated in
membrane insertion and pore formation. There is a chymotrypsin
cleavage site in this loop that is required for toxicity. Domain 3
has a hydrophobic patch thought to be involved in protein-protein
interactions. Domain 4 appears to be a separate domain and shows
limited contact with the other three domains; it would swing out
of the way during membrane insertion. It is required for binding
to the receptor; the small loop is involved in receptor
recognition.
[23] PTM: Proteolytic activation by furin or a furin-like protease
cleaves the protein in two parts, PA-20 and PA-63; the latter is
the mature protein. The cleavage occurs at the cell surface and
probably in the serum of infected animals as well; both naive and
cleaved PA are able to bind to the cell receptor. The release of
PA20 from the remaining receptor-bound PA63 exposes the binding
site for EF and LF, and promotes oligomerization and
internalization of the protein.
[24] MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were
performed that showed that the residues present in the small loop
of domain 4, and not the ones in the large loop, are involved in
receptor recognition.
[25] SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02; Search time 8.2539 Seconds
(without alignments)
4636.784 Million cell updates/sec

Title: US-09-848-909A-11
Perfect score: 3774
Sequence: 1 EVKQENRLINESSESSQGLL.....TSTNGIKILIFSKGVGIG 735

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	3745	99.2	764	1 PAG_BACAN
2	235.5	6.2	204	1 YPSI_BACAN
3	190.5	5.0	1630	1 MSP1_PLAFK
4	190.5	5.0	1639	1 MSP1_PLAFW
5	174.5	4.6	979	1 P115_MYCHR
6	173	4.6	1658	1 YM67_YEAST
7	171	4.5	1000	1 S155_YEAST
8	170.5	4.5	1230	1 SMC3_YEAST
9	164.5	4.4	2334	1 WAPA_BACSU
10	163	4.3	1385	1 FATI_SCHPO
11	162	4.3	1276	1 BXD_CLOBO
12	160	4.2	1803	1 YJL3_YEAST
13	159	4.2	1420	1 SRB3_YEAST
14	158.5	4.2	1127	1 CAG3_HELPJ
15	158	4.2	1957	1 SF02_SCHPO
16	157.5	4.2	2116	1 MYS2_DICDI
17	157	4.2	1790	1 US01_YEAST
18	156.5	4.1	1487	1 MDS3_YEAST
19	156	4.1	2867	1 RBP2_PLAVB
20	155.5	4.1	1208	1 PCP1_SCHPO
21	155.5	4.1	1460	1 N159_YEAST
22	155	4.1	938	1 YMJ3_CABEL
23	154	4.1	1029	1 RIP3_RAT
24	153.5	4.1	1233	1 YF16_YEAST
25	152.5	4.0	1024	1 RIP3_MOUSE
26	152.5	4.0	1116	1 YK54_AQUAE
27	152.5	4.0	1271	1 Y338_MYCGE
28	152.5	4.0	2869	1 RBP1_PLAVB
29	152	4.0	1882	1 Y460_MYCPN
30	151.5	4.0	1397	1 CID_DROME
31	150.5	4.0	1358	1 SIR4_YEAST
32	150.5	4.0	1628	1 NAGH_CLOPE
33	150	4.0	1037	1 KCC4_YEAST

ALIGNMENTS				
RESULT 1				
PAG_BACAN				
ID	PAG_BACAN	STANDARD;	PRT;	764 AA.
AC	P13422; Q9FSR7; Q9RH69; Q9RQ2;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins			
DE	translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].			
GN	PAGA OR PAG OR PXOI-110.			
OS	Bacillus anthracis.			
OC	Plasmid pXOI.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1392;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89172073; PubMed=3148491;			
RA	Welkos S.D., Lowe J.R., Eden-McCutchan F., Vockin M., Leppla S.H.,			
RA	Schmidt J.J.;			
RT	"Sequence and analysis of the DNA encoding protective antigen of			
RT	Bacillus anthracis.";			
RL	Gene 69:287-300(1988).			
PN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=28, 33, BA1024, and BA1035;			
RX	MEDLINE=92214082; PubMed=10197996;			
RA	Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;			
RT	"Genetic diversity in the protective antigen gene of Bacillus			
RT	anthracis.";			
RL	J. Bacteriol. 181:2358-2362(1999).			
PN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=V770-NP1-R / ATCC 14185;			
RX	MEDLINE=20359347; PubMed=10899854;			
RA	Cohen S., Mendelson I., Altboum Z., Kobiler D., Elhanany E., Bino T.,			
RA	Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M.,			
RA	Kronman C., Velan B., Shaffer A.;			
RT	"Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus			
RT	anthracis spore vaccines protect against anthrax.";			
RL	Infect. Immun. 68:4549-4558(2000).			
PN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sterner;			
RX	MEDLINE=99445483; PubMed=10515943;			
RA	Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,			
RA	Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,			
RA	Martinez Y., Rieke D., Svensson R., Jackson P.J.;			
RT	"Sequence and organization of pXOI, the large Bacillus anthracis			
RT	plasmid harboring the Anthrax toxin genes.";			
RL	J. Bacteriol. 181:6509-6515(1999).			
PN	[5]			
RP	DOMAINS.			
RX	MEDLINE=91332080; PubMed=1651334;			
RA	Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;			
RA	"The carboxyl-terminal end of protective antigen is required for			
RT	receptor binding and anthrax toxin activity.";			

P53968 saccharomyc
O14157 schizosacch
P40468 saccharomyc
P13611 homo sapien
Q00496 clostridium
Q92KWS helicobacte
P16466 proteus mir
P09959 saccharomyc
P10844 clostridium
Q03213 saccharomyc
Q49413 mycoplasma
P18494 saccharomyc

Db 1051 NTKN-----TWPNSTLISELTSKEDSKNSITNSSKSDIET 1087

RESULT 15

T31105

hypothetical protein 2 - Haemophilus ducreyi

C:Species: Haemophilus ducreyi

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T31105

R:Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.

J. Bacteriol. 180, 6013-6022, 1998

A:Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.

A:Reference number: Z20984; MUID:99030326; PMID:9811662

A:Accession: T31105

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4919 <WAP>

A:Cross-references: EMBL:AF057696; NID:G3929021; PID:G3929023; PIDN:AAC79761.1

C:Genetics:

A:Gene: lapA2

Query Match 4.7%; Score 178; DB 2; Length 4919;

Best Local Similarity 20.0%; Pred. No. 2.4;

Matches 182; Conservative 126; Mismatches 328; Indels 274; Gaps 41;

QY 6 NLLNESSSSQGLGYFSDLNFOAPMVVTSSTGDL-SIPSSSELENIPSENQYF---Q 61

Db 154 NNSIKEGNSQLVLLGEN-KNLGSOAAKTIHQVIGDOESKISGGLGVFGEKADLIINP 212

QY 62 SAIWSGFIVKKSDEYTFATS-----DFKLYTDSQNK-----KEV 145

Db 213 NGVTLNGVKTINTDRFVASTSEVVEPHIKQLNVQKGIIGKGVATNGLSHPDVVAKNI 272

QY 83 -----ADNHTVMVDQDVINKASNSNKIRLEKGLYQIKI-- 118

Db 273 EQQKGVSIKEDSKPAKLVNTPAAGNLTYDVNTRDV-NRNTNPKPITDTRKDNIAISG 331

QY 119 -----CYQREN-----TEKGL----- 391

Db 332 ESAGSMYGRNIFVTRKAGVNHQGVIFAEDDIIILTDGNSRLNKVYADYRVVVGKDI 391

QY 146 ISSDNLQLPELKQKSNRSKRKSTAGPTVPDRDNDGIPD-SLEVEGYTVYVKNKRTFLS 204

Db 392 ELANNQIHAQQIILNATHGVKLDNGSSVSNNNLIGISALNLTLENATVSANN---LS 447

QY 205 PWISNTHKKGKLTYSSEKWSSTADPYSD-----FEKVTGRIDKNVSPPEARHP 254

Db 448 FRYTNDTKLNLKSVS-----ARAADLQSGNLNLDKASVLAHKULTNISNDVSLNQSK 501

QY 255 LVAAYPVHVDMENIILSKNEQSTONTDSETRTISKNTSTRTSTSEVHGNAEVHSAFF 314

Db 502 LSA-----NNLAKIKVRLNLSNLSANNLTNTSNNITLKNK-----SKP 543

QY 315 DIGGSVSAFSSNSSTVAIDHSLSLAGERTWAETWGLNTADTARLNANIRVYVGTGAPI 374

Db 544 TAGNMT---LNVTVNVTLNDSELA-ANNLTNLVTKNVTLNDSKLSANKLDLV----- 594

QY 375 YNVLPITSLVLGKQNLATIDADENLSOI-----LAPNNYPSKNLAPALNAQKFF 427

Db 595 -----TDNVTLSKSTLSAGELTFKKVKVTVLNDSELAANN-----LSLNA----- 636

QY 428 SSTPITMYNQFLELEKT--KQLRL---DTQVYGNIAFYNPENGRVVDGTSNWSVLP 482

Db 637 -SHNVTLNKKSKLSAQKADIKAVNLTLNDTTE---LTAKNLDINSTTITNNGTIAGIFA 691

QY 483 QIOETTARLIFNGKDLNVERRI-AAVNPSDPLETTKPDMTLKEALKAPGENEENGNIQ 541

Db 692 NI---TTEKLNKKEKALILAEQNLNFTVNGSH--YENKGDIVSKDRTVTFSKNS----- 741

QY 542 YQKGDITEFDNFDQOSTQNIKNQLAELNATNI-----YTVLDKIKLNK----- 586

Db 742 -----DFTSGSKLVNAQNL-KVNVNFTISQGGDITLIGNVTLNASGTTFTNSGN 791

QY 587 ---MNLIRDKRFHYDRNNIAVGAD---ESVVK-----EAHREVINS-- 622

Db 792 LTTVKTLVDYGDIQNFTNKGNLTVGEDLHKSKTKITNDGKLISIKNLNISSEADFINNGT 851

QY 623 -----STEGLLNIDKQIRKILSGYIVIEDETEGLKEVINDRYDM---LNIS-- 667

Db 852 LLGIEALKIATKGNFTNKEKAI--LASNGLDLSVAEGKKTFFNNGTIESGKLNLTNTGA 909

QY 668 -LRQDQKTFIDFKKYNDKLPYISN-----PNYKVNVAVTKENTTINPSENGDTSSTNGI 721

Db 910 FLNVDMATRSFGVLNITSGVSNNGTLISHERLN---ITSAANFTNES-NGTVMSNGL 965

QY 722 KKILIFSKKG 731

Db 966 LNII--AKQG 973

Search completed: May 3, 2004, 19:41:19

Job time : 16.0855 secs

A;Cross-references: GB:AB05176; PID:gl2724625; PIDN:AAK05715.1; GSPDB:GN00146
A;Experimental source: strain IL403
C;Genetics:
A;Gene: YqfG

Query Match 4.8%; Score 179.5; DB 2; Length 1072;
Best Local Similarity 19.4%; Pred. No. 0.21;
Matches 160; Conservative 142; Mismatches 342; Indels 179; Gaps 30;

```
QY 2 VQENELLNSESQGLLGYVFDLNFQAPMVVTSSTTGDLSIPSSLEN----- 52
DB 300 VEQDVVASSSTODANASLYPISEASSVTDNLNLSISLSDSSISSTOTENSQSGASTA 359
QY 53 ---IPSENQYFQSA---IWSGFIKVKSDSEYTFATSDADNHVTVMDQEVINKASNSNKI 106
DB 360 EISYDSENSNLSSNQINSNSSEKDSQSSGLSSMSNEESHSNSNINETNSSEI 419
QY 107 R-----LEKRLYQIKIQRENPTKGLDFKL----YWTDSQNKKEVSSNLDLPKL 157
DB 420 TNLPPSPNTENSVDQTSSEASTNSNSISLSPSNSISSTSDSESATNSDFSNAVA 479
QY 158 QKSSNRKKRSTAGTVPDRNDGI---POSLEVEGYTVQVAKRTPFLSPWISNIHEK 213
DB 480 NNSLASVNNSSSVLSTSTADNLGINQSGSNLTKD--SSSISTSGAFLS---SNQTS 534
QY 214 KGLTKYKS-----TADPVSDFPEKVTGRDKNVSPPEARPLVAAPV 261
DB 535 EASTNSNSISLSPSNSISSTVLESTSSNSFNVAEAVNGLASVNNSSSVLSTSTA 594
QY 262 VHVDM-----ENT-----ILSKNEDOSTONTSETRTISKTSTRTHTSEVH 304
DB 595 DNEINQFGSDNLTKDSSEISLSTSGAFLSSNQTSSEASSNSMSINSPLSLSTNSESA 654
QY 305 GNAEVHASFPDGGSVSAGFNSNSSTVAIDHSLSL--AGERTWAEWTGLNTADTARLNA 362
DB 655 TNQ-----SNSSEATKVDNNSST---HSSNINLGSNDSDSDSDSDSLSNL-- 699
QY 363 NIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENOLSQ---ILAPNYYPSKNLAP 418
DB 700 -----SSPNLETQTISSKSEFVNINSEPKVSSNSVQENSTDE 742
QY 419 IALNAQKFFSTPTMYNQVFLKTKQLRLDQVYGNIAFYN---FENGVRVDTGS 475
DB 743 MSTNPKSSISPTSTSSSQKESQN---LLNTTEGINNPTFNNSSSENSAASILT-- 797
QY 476 NWSEVLQIOETTARIIF-----NGKDL-----NLVERRIAANVPSDPLETTKPDMT 522
DB 798 SYSNNSESSETGCLYISNEAQDNGSEISHSPLSPSNSENNVSISQSAILESSKSTN 857
QY 523 LK-EALKIATGFNEP-NGNLQYQKDIITEPDFDQDTQSONIKXQLAELNATNIYVLDK 580
DB 858 KRSSLSIINSTSHPNQEDNQSSD-----EVKSNNVBSILQLNSISNKTWMS 909
QY 581 I---KLNAKNILIRDKRFHYDRNNTAVGADESVEKAEHREVNINSSTEGLLNIDKIRK 637
DB 910 LTSQKLGVITLPSKSKVTNEKNENSTVSEKLIKTPQCN-DESNLGOITALDLSFNK 968
QY 638 ILSGYVIEIEDTEGLKEVIN-----DRYDMLNISSLRQD-GKTFIDFPKYNKPLPY 688
DB 969 -----EVETMEDSKTVPKVLNENGRSQNNKTSTIAKDKNKVFPKSEFNSKIL-- 1019
QY 689 ISNPYKVNVAVTKENTINPENGDTSTNGIKKILIPSKG 731
DB 1020 -----DSDNNILKTKVLLKKG 1036
```

RESULT 14

JC6009
surface-located membrane protein lmp3 precursor - Mycoplasma hominis
C;Species: Mycoplasma hominis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: JC6009
R;Ladefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.

J. Bacteriol. 178, 2775-2784, 1996
A;Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene sytse
A;Reference number: JC6009; MUID:96213016; PMID:8631664
A;Accession: JC6009
A;Molecule type: DNA
A;Residues: 1-1302 <LAD>
A;Cross-references: EMBL:X95601; NID:gl197335; PIDN:CAA64858.1; PID:gl197336
C;Genetics:
A;Gene: lmp3

A;Genetic code: SGC3
C;Superfamily: surface-located membrane protein lmp3; tetratricopeptide repeat homology
C;Keywords: duplication; membrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-1302/Product: surface-located membrane protein lmp3 #status predicted <MAT>
F;957-992/Domain: tetratricopeptide repeat homology <TT1>
F;993-1026/Domain: tetratricopeptide repeat homology <TT2>
F;1089-1120/Domain: tetratricopeptide repeat homology <TT3>
F;1154-1190/Domain: tetratricopeptide repeat homology <TT4>

Query Match 4.7%; Score 178; DB 1; Length 1302;
Best Local Similarity 18.0%; Pred. No. 0.33;
Matches 138; Conservative 141; Mismatches 285; Indels 202; Gaps 28;

```
QY 1 EVKQENELLNSESQGLLGYVFDLNFQAPMVVTSSTTGDLSIPSSLENIPSENOYF 60
DB 476 QVDEANKSIKEQLNALIDKANTLLPOLNDSSEIVAKESLNAEITNANKAVNQNDASM 535
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTVMDQEVINKASNSNKIRLEKGR--LYOIKI 118
DB 536 QSA-----KSSLDKVKTIQNLQTEFNKDKAKFKELEQTRKDIDNFLT 579
QY 119 QYQRENTEKGL-----DFKLYWTDSONKKEVSSDNLQLPELKQKSSNR--KRS 168
DB 580 DDVKNNNYATLVKDLTNAKDKSVTKSNKSEIIAAND-----ELKQALDKAKVADQI 635
QY 169 TSAGTVPDRNDGIPDSLEVEGYTV---DVKNKRTFLSPWISNIHEKGL----- 216
DB 636 DEANKSIKEQLSDSITNANQLLNKLVDSKDIOAKATELSQEQSASQELNLANPTSMQS 695
QY 217 -----IKYKSSPEKWTASD-PYSDEPKVTGRDKNVSPPEARPLVAAPVIVHDM 266
DB 696 AKESLDKAVTEITKKLETFNKKDKVDFKLEKTRKDIDEFINTKNTNP----- 743
QY 267 ENILSKNEDOSTONTSETRTISKTSTSTRTHTSEVHGNAEVHASFPDGGSVSAGFSN 326
DB 744 -----NYSTLISELTSKSDSKSNVTSSNKSIDI----- 771
QY 327 SNSSTVAIDHSLSLAG-ERTWAEWTGLNTADTAEALNANIRYVNTGTAPIYVNLPTTSLVL 385
DB 772 -ETANTELKQALAKANTDKAQADNLAKSTKE--QLNNSISSANTLLAKLTDKONTI---- 824
QY 386 GKNOTLATIDADENQLSQILAPNYYPSKNLAPIALNAQ-----KFFSSTPTITMYNQFL 440
DB 825 --QQAQTELEKEVQKAVASNN-TASMQSAKSLDAKVTETKKLE---TFNKDKDV 877
QY 441 ELEKTKQLRLDQVYGNIAFYNFENGVRVDTGSKNWSSEVLQIOETARIIIFNGKDLNL 500
DB 878 KFKELEQTRKDIDEF-----INTKNTPDYSTLISELTSK----- 912
QY 501 VERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPQNLQYQKDKITEPDFNEDQOTSQ 560
DB 913 RDSKNSITNSNKSIDIETANTELQALAKA-----NTDKQAD-NLARSTKE 958
QY 561 NIKQQLAELNATNIYVLDKIKLNAKNILIRDKRFHYDRNNTAVGAD---ESVVKBAHR 617
DB 959 QLNKSISAN-----TLAKLT-----DKDNTTQQAQTELEKEVQKAVQ 997
QY 618 EVINSSREGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDR-YDMLNISSLRQDGKTFI 676
DB 998 AVAGNNTASM-----QSAKSLDAKVTETI--TKLETFNKDKVDFKLEQTRKDIDEFI 1050
QY 677 DFKYNDKPLPLYISNPYKVNVAVT-----KENTINPSENGDTST 718
```



```

385 QY LGKN-QTLATIDAD-ENQLSQI-----LAPN-----NYP-----SKNLAPALNAQCKF 427
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
755 DB IDKSIETLTTEIKNGSENHIDEIKGQIDKLKVPKNTMFNEPKEIEKIEINIVKIDKKK 814
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
428 QY -----STPITWNYQFLEKTKQLRLDITDQVYGNATYFNFENGVRVDTGSGNWSE 479
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
815 DB NIYKEIDKLLNEISKIENDKTSLEKLNINLSYGVKSLGNLFQQIDEEKKAAEHTIKAME 874
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
480 QY -----LPQIQETARIIFNGKOLN-----VERRIAANPSDP-----LETT-----K 518
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
875 DB AYVIDLDINIKKSQEI-----EKEMNNIMDKMIDHKEMKALNISHDDYKIYTTSKNHEEK 931
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
519 QY PDMTKEALKIAFGFNEP---NGNLQYQGGKITFDFNFDQOTSONIKQLAELNATNIY 575
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
932 DB ISDIRKNSIKIQDFSEESYINDIKKELEKNVLE-----SQNNVTIDINQYLSKIE--NIY 984
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
576 QY TVLIDKIKLAKNMLIRDKRF--HYDRNNIAGADESVVKEAHRVINSSTEGLLNIDK 633
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
985 DB NIL--KLNKIKIIDKVKYEYDTEIEKN-----KKINAELS-----1020
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
634 QY DIRKILSGVIVETDETEGKE-----VINDRY-----DMLNSSLRQDGKTDFDKYND 683
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1021 DB -EKI-----ITQKENSLSKECOSKIKSTIDDNYVSECINITNL-----KTVIVNEKN- 1069
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
684 QY KLPYIYISN-PNYKVV-----YAVTKENTIINPSENGDTSN-GIKKILIFSKK 730
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1070 DB -INTYFKNAEYQNVSLNPNFIEMADTKSQVILNIKNGNTNTDYNIKELKEHKKK 1126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
B89921
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89921
R:Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc-
ma, A.; Mizutani-Jui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11416146
A:Accession: B89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <R>
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BAB42527.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebha

```

[illegible]

286 QY TRI-----SKNTSIRTHTEVHGNAEVBHAFDFDGGVSA-----GFSNSNSSTVAIDHS 337
2720 Db LATAKQQAQDALLAQMTLSDAQKQS-----ITQDLSATQVGTQVQSVKQATNLDNA 2771
338 QY L-----SLAGERTWAETMGLNTADTARLNAHRYNTGTAPYINVLPITS-----LV 384
2772 Db MQLRNSIAKNEVKASQPVADATDTKQNA-----YNTAVTSAENIYNATSQPLDPSAVT 2827
385 QY LGKNOITLATIDADENOL--SOILAPNYYPSKNLAPIA--LNAOKKFFSPTIMTNYN-QF 439
2828 Db QAAQ-----VNTKRTALNGAQNLANKKOFTAINRLSHLNAQKODLNTQVTPAPNIST 2883
440 QY LELEKTKQLRL-----DTQV-----YGNIAI-----YFEN 466
2884 Db VNOVYTKAEQLDOAMERLINGIQDKQVQKQSVNFTDADPEKQYANNVATAENIINQAN 2943
467 QY GRVRVDGTSNWSVEVLPIQIE--TTARIIFNCKDLNLVERRI--AAVNPSPDLET-----516
2944 Db G-----TNANQSQVEAALSTVTTTKQALNG-----DKRYTDKNNANQITLSTLDLNN 2999
517 QY -----TKPDMTILKEALKIAGFNEPNGLQ--YQGHDIITFDNFDP-----555
2993 Db QKGAVTGNINQAHTVAEVI--QAOTQAEIANTMGNLKNLNDKDTTLGSONFADADPEK 3050
556 QY -----QQTSONIKQOLAEINATYIVLDKIKLAKMKMILLRDKRFHYDRNNIAVGADSV 611
3051 Db KNAYNEAVRAENTILNKSTGNV--PKQVEAAMN-----QVNTYKALNGTQNL 3098
612 QY VKEAHREVINSSTEGE--LLNIDKDIRKILSGYIEIEDTEGLKEVINDRYDMLNIS-- 667
3099 Db --EKAKQHANTAIDGLSHLTNAQKEALKQLVQOQTTVAEAGQBOKAN-----NVDAAM 3150
668 QY --LQ-----DGKTFIDFKYNDKPLIYSNPYKYNVY--AVTKENTILINSEN 713
3151 Db DKLRQSTADNATTKONQYTD-----ASPEN--KKDAYNNVTTAAGIIDQTTN 3196

RESULT 11

G90599
Hypothetical protein MYPU 7030 [imported] - Mycoplasma pulmonis (strain UAB CTB
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: G90599
R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; M
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycop
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: G90599
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-752 <KUR>
A:Cross-references: GB:AL445566; PID:gi4090118; PIDN:CAC13876.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU_7030
A:Genetic code: SGC3

Query Match 4.8%; Score 181; DB 2; Length 752;
Best Local Similarity 19.5%; Pred. No. 0.1;
Matches 161; Conservative 129; Mismatches 269; Indels 268; Gaps 43

QY 41 GDLSTPSSELENIIPSENQYFQSAIWSGFIKKVKSDEYTF-----ATSAHNHTVMYVD 92
Db 20 GTISGISYISQLPNESAL-----VKRAMELPLNKNQSVLNKSIHELINQWQD 69
QY 93 DOEVINKASN-----SNKIRLEKGR--LYQI---KIYQRENPTXKGLDFKLYWTDQ 140
Db 70 ENKVSNI SANDFNKVFQSKTTPLENGEKITYSVLGKDIYFQVNPQKTIKSVKITSK 129
QY 141 NKEVITSSNLQIPELKQKSN--SKKRSKTAGTVPDRDND-----GIPDSLEVBGYTV 194
Db 130 ISKDVYMD-----KQRLNDFAKLRVNFKGSASKEQSDIWAQNFQDKSKLEIKY 182

3

QY 577 VLDKIKLAKOM--ILIRKRPYDRNN-----IAGV-ADBSVKEAAREVINGSTGLL 628
Db 1466 IQDKLADFKONNFVGIADLSTIDYHNLLTKFLSTGWFENLAKTVLSNLDGNLQO-M 1524
QY 629 LNKIDK--IRKILSGYIIEIDTEGLKEVINDRYDMLNISLROQKTFIDFKYNDKLP 686
Db 1525 LNISQHCVKK-----QCPNSGCFRHLDE-----RECKCLLYKQEGDKC- 1566
QY 687 LVIISPNKYVYATKENTINPENG-----DSTNGIKKI 724
Db 1567 --VENPNTCN-----EN-----NGCCDAKCTEEDSGNG-KKI 1599
RESULT 8
B64635
toxin-like outer membrane protein HP0922 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: B64635
R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 368, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: B64635
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2529 <TOM>
A:Cross-references: GB:AE000602; GB:AE000511; NID:92314060; PIDN:AAD07969.1; PID:g231406
Query Match 4.9%; Score 185.5; DB 2; Length 2529;
Best Local Similarity 21.0%; Pred. No. 0.37;
Matches 158; Conservative 94; Mismatches 261; Indels 241; Gaps 36;
QY 123 ENTEGLDFKLYWTSQNK---KEVISSDNLQLPELKOKSSNRKXSTSGAFTVDP 178
Db 729 QNPAPASVSGVWL--QNKTSYNGIYID-----PNLSGQSGQNTLSTYANLF--- 779
QY 179 DNDGIPDSLEVEGYVDVKNKRTFLSP---WT-----SNIHEKKGLTKYK 220
Db 780 ---GRGFSVNIQNTLIGNTSVNSGLIWHGGFGYITGTFSAANIY---LTNNF 832
QY 221 SPSKYST-----ASD-----PYSDFEKTGTGIDKNVSPEARHPLVAAYPIV 262
Db 833 KTGEGVNSDGGGANITFKASDNIWDLNDAETVTMIOGAS---OHSYATFDALN 889
QY 263 HVDMEIILSKNEDQSTQNTDSETRTISKNTSTSRHTTSEVHGNAEVHASFDDIGSVSA 322
Db 890 NISVTNSSFS-----DMTWGKFSFAKNISFS-----NASF-----S 921
QY 323 GFSNSNSSTVA---IDHSLAGERWAEWTMGLNTADTARLANIRYVTGTAPIYVNL- 378
Db 922 GFTNPGSSVISANATNSLSFINSR-----LNGGAVYNLQANSIFPN-NTQAVFNVL 973
QY 379 -----PTTSVLGNQCLATIDADENQLSOILAPNPNYPSKNLAPIAL--NAOKFES 428
Db 974 SRGTSFNATQLLNTNFTLSS-----QSLNFGDTTLQNNANITLGNKSQAAPK 1025
QY 429 STPTMNYNPFLEKTKQLRLDQVYGNATYFNPNFNGRVRVDTGNSWSEVLPOIQTET 488
Db 1026 NS-LTLDNNSLSDNQSVLNANNTSAFNQASNLINYGS-----QATF 1068
QY 489 ARIIFNCKDNL-VERRIAAVNPSPLETKPMTLKEALKIAGFNEPNGLNVOYQKDI 547
Db 1069 NSLFFNGGTLNASKSKNANASAFSNNTT---INLDSVLSASNTSSLNANINFGASQ 1125
QY 548 TEF-----DFNPDQQTSONIKNOLA-----E 568
Db 1126 ADPGGNTIIDTASPNFDSAGSLNPNLNTANGALNFNGYTPSLTKALMSVSGQFVLGNNG 1185
QY 569 LNATNIVTLDKIKLAKMILLIDKEF-----HYDRNNIAV 605

Db 1186 INLSDI-NIFDNTKTSVTYNILNAQKGTIGISGANGYEKILFYGMKIQNATYSDNNNIQT 1244
QY 606 GA-----DESVVKEAHR-----EVIN--SSTEGLLNIDKDI-----RKILSGYIVE 645
Db 1245 WSPINPLNSQIIQESIKNGDLTIEVINPNPNSASNTIFNIAPELYNYQASKNQPTQYSD 1304
QY 646 IEDTEGLKEVINDRYDMLNISL-----RDQKTFIDFKYNDKLPFLY---ISPNPY 694
Db 1305 YSDNQA-----GTYYLTNSNIKGLFTPKGSQTPQAPQYSPFPNOPLSSLIYNNKGFSSENL 1359
QY 695 KVVYAVTKENTIN--PSENGDTSTNGIKKI 725
Db 1360 KTLGILSQNSATLKEMIESNQLDNTN-INEVL 1392
RESULT 9
T28576
rhostry protein - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28576; A45521
R:Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem. Parasitol. 76, 329-332, 1996
A:Title: Comparison of two members of a multigene family coding for high-molecular mass
A:Reference number: Z20507; MUID:97077455; PMID:8920022
A:Accession: T28576
A:Status: preliminary; translated from GB/EMBL/DBR
A:Molecule type: DNA
A:Residues: 1-2401 <SIN>
A:Cross-references: EMBL:U36927; NID:91041784; PID:g1041785; PIDN:AAB41263.1
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A:Title: Identification of the gene for a Plasmodium yoelii rhostry protein. Multiple c
A:Reference number: A45521; MUID:91101660; PMID:2270106
A:Accession: A45521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 2260-2401 <XEE>
A:Cross-references: GB:M34281
Query Match 4.9%; Score 184.5; DB 2; Length 2401;
Best Local Similarity 21.6%; Pred. No. 0.38;
Matches 181; Conservative 120; Mismatches 316; Indels 221; Gaps 43;
QY 1 EVKQENLLASESSSQGLGYFSDLN-----FOAPMVVTSSTTGDLSI 45
Db 402 EYAKENVOLNRYKSNILEIKKHNDQINIDNIKEAKQNYDQFKEHKTIPPNEKYQK 461
QY 46 PSSELENIPSENOYFQSAI--WSGFIVK---KKSDEYTFATSADNHYTMVWDDQEV---I 97
Db 462 PSIBIKIMKDB---FLSKVKNKYNDFKVYKEKVESEHNKFTLTNKTKEVSDEIKKYE 518
QY 98 NKASNSKIRLEKGLYQIKIQORENPTEKGLDFKLTWDSQNKKEVISEDNLQLPK 157
Db 519 NKFNDSKSLNETKK--SIEEBYQNTLKKVDD---YIKVCLNTNELINTCHNQKQTLK 573
QY 158 QKSNNSRK--KRSTSAGTVPDRDNDGIPD-----SLEVEGYVDVKNKRTFLSP 205
Db 574 DKLQNKIKETNSIDKIYTDKFNILTDKKELETKFTGLSLNNHESNNKELLTYFD 633
QY 206 WISNHEKGLTKYKSPKWKSTASDPSDEKVTGIDKNVSPEARHPLVAAYPIVHVD 265
Db 634 LKANLGNKNENLYKQFNEK---EKAVEDIKKNVDINKIVSN-----IEIT 677
QY 266 MENILSKNEDQSTQNTDSETRTISKNTSTSRHTTSEVHGNAEVHASFDDIGSVSAGFS 325
Db 678 IYTSIYNED-----TENE---IGKSIELLNTKVL-----KVKANVT 713
QY 326 NSNGSTVAI--DHSLSLAGERTWAEWTMGLNTADTARLANIRYVTGTAPIYVNLPTSLV 384
Db 714 NLNEIKELKDYDFQDFGK-----EKNIKYPDEN--KIKNDITLQNK 754

A:Cross-references: GB:AL44566; PID:gl4090108; PIDN:CAC13866.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MIPU_6930
A:Genetic code: SGC3

Query Match 5.1%; Score 191.5; DB 2; Length 1125;
Best Local Similarity 19.2%; Pred. No. 0.05;
Matches 166; Conservative 163; Mismatches 328; Indels 207; Gaps 40;
QY 3 KOENLLNESSESSQGLLYGYSDFNFOAPMVVTSSTGDLSPSSSELENIPSENOYFQS 62
DB 120 ENDNSVNNKNSK-----NDN-----LTLTKVRLGHVNV-LNOSGNPTKLAISK 168
QY 63 AIWS-----GPIKV-----KXSDYTFATSDADNHVTMWVDDQE-----VI 97
DB 169 VILHNKLDVLGLTEITNENGVTIINELNKSSDKWNVIVSLKRGTTGSSGQBEHVGI 228
QY 98 NKASNSKRLP-----KRLVQIKIOYORENPTKGL----- 130
DB 229 YK-----ENKLTLSFDDKESKGFYENKLW-----DPPFKGQKIDVRPPFGVKSTKGN 282
QY 131 --DFKLYW--TDSQNKKE---VSSDNLQLPKLSNSRKRSTSGTPTDPRDNGI 183
DB 283 KNDFTVVFHSDAPGVKEERGEISAKYSGQGHKEVAEALRTKEVMEYFDSI-----DGV 337
QY 184 PDLVEGYT-----VVKKRTTSLP-W--ISNIHEK---K 214
DB 338 NNELPFMGDTNKLNGEAKAFPLQSGYKSIUKDVKENATSLAORWGEYANHYDKIYK 397
QY 215 GLTKYKSSP--EKWSTASDPYSDFEKTGRIDK-NVSPPEARP-----LVAAYPI 261
DB 398 GDJLVKNSGFYDLWKYFDDNLLNKEFKTKVEEYRSSKKAKYKGEYSYVLHAI SDHTI 457
QY 262 VHYDMENIILSKNEDOSTONTSETRTISKVTSSTHSTSEVGHNAEVAHFDDIGGSVS 321
DB 458 VYTDL--ILLQKOD---QNKSENK--DENSDNSKQNDKPKTNNEQONTQDSSKKIS 509
QY 322 AGFSNSNSVAIDHSLAGERTWAEITGLTADTARLNANIRYVNTGTAPIYVLPPT 381
DB 510 DAGONNSNTNEKQKLDSDQE---SKNAIKSQNDQKDSNLSSKNTQPSKSSPOI 566
QY 382 SLVLGKQNTLATIDADENQSLAPNPNYPSKXLAPIALNAQKFSPTITWYNQFLE 441
DB 567 NPNLENNQETSHSNGENDSDSKQNTSNGRTKN-----DLRSEKQNLK----- 611
QY 442 LEKTKQLRLDQVYGNIAVNFENGVR---VDTGSNWSEVLPOIQTETARIIFNGKDL 498
DB 612 ---TKNPSNSNVETKNETQNNNSSTKXDEIDTSA-----KTQDSTNSLNKEEKT 661
QY 499 NLVERRI-----AAVNPSPDLETTKPDWTLKEALKI--AFGNFENGNLQVQKDI TE 549
DB 662 NOVETKNTBSNNSNSTKQENSSTKKEBISKSESNVANSNSTKQENIDNKKEBISK 721
QY 550 PDRFDQOTSONIKNO-LAEINATNIYTVLDKIKLAKNN-----TLIRDKRPHYDRNMI 603
DB 722 SESVNNNSNTQNTQETPTESQNNVILGKPNKQSLQNAIDVSAKKVIGYWNINE 781
QY 604 AVGADESVAKEAHREVINSTEGLL-----LNIDKIRKILSGYVIEIETEGEKEVINDR 659
DB 782 SVGKSSASKAFVAKVIDHNKLDIVGIGLVHEETLTKI-----VEEMNKLSDSSDK 834
QY 660 YDMLNISLRDQGTTP-IDFKYNDKLPYISNPNKYVNTVATKENT---IINPSENGD 715
DB 835 W--VQVISEKKQGGFFVNLARYTG--VIVKEN---KFNIESFKNQNTNGHLYENQPNW 897
QY 716 TSTNGIKKI-----LIFSCKG 731
DB 888 SSFNTSEKVSVPFPFGIKFSTKG 911

RESULT 7
S05603

major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (s
N; Alternate names: gp195 surface antigen
C; Species: Plasmodium falciparum
C; Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jun-2000
C; Accession: S05603; S04850
R; Myler, P. J.
Submitted to the EMBL Data Library, April 1989
A; Reference number: S05603
A; Accession: S05603
A; Molecule type: mRNA
A; Residues: 1-1639 <MYL>
A; Cross-references: EMBL:X15063; NID:g9896; PIDN:CAA33163.1; PID:g9897
R; Myler, P. J.
Nucleic Acids Res. 17, 5401, 1989
A; Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plas
A; Reference number: S04850; MUID:89345116; PMID:2668887
A; Accession: S04850
A; Molecule type: mRNA
A; Residues: 1504-1639 <MYL2>
A; Cross-references: EMBL:X15063
C; Superfamily: major merozoite surface antigen
C; Keywords: glycoprotein; merozoite; surface antigen
F; 1-19/Domain: signal sequence #status predicted <SIG>
F; 20-1639/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 5.0%; Score 190.5; DB 2; Length 1639;
Best Local Similarity 20.6%; Pred. No. 0.11;
Matches 171; Conservative 138; Mismatches 255; Indels 265; Gaps 47;
QY 5 ENRL-LNESSSSQGLLYGYSDFNFOAPMVVTSSTGDLSPSSSELENIPSENOYFQS 63
DB 927 ENILSGKNKIYQELIGQKSE-NF-----YEKILKDSDFYNE 965
QY 64 IWSGFIVKKSDEYTFATSDADNHVTMWVDDQEVINKASNS-----NKIRLEKGRLY- 114
DB 966 SFTNFVSKADD-----INSLNDESKRKLLEEDINKTKTLQSLFDLYNKYKILKLERLD 1020
QY 115 -----QIKYOQRENTKGLDFKLYWTSQNKKEVISSDNLQLPKLSNSRKRSTSGTPTDPRDNGI 168
DB 1021 KKTVGVYKMQIKKLTLLKEQLESKL--NSLNPKHVL--QNFVFNKKEKEAIEAETEN 1076
QY 169 TSAGTVPDRDNDGIPDSLEVEGYTVDKNRTFLSPWISNIHEKKGLTKY---KSSP-- 223
DB 1077 T-----LENTKILLKHY-----KGLVKYANGESSPLK 1103
QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPPEAR-----HPLVAAYPIVHVD 265
DB 1104 TLESISQTDENYASLENFKVLSKLEKQKDNLNLEKKLSYSSGLHLLIA-----E 1156
QY 266 MENIILSKNEDQSTONTSETRTISKVTSSTHSTSEVGHNAEVAHFDDIGGSVSAGFS 325
DB 1157 LKEVKKKN---YTCNSPSENNT-----DVNNALESYKKEFLPEGTDAVTVS 1200
QY 326 NSNSTVAIDHSLAGERTWAEITGLTADTARLNANIRYVNTGTAPIYVLPPT----- 381
DB 1201 ESGSTLESQSPKPKPASTHVGAESE---NTITTSQ--NVDDEVDVVIIVPIFGESEEDYDL 1256
QY 382 -SLVLGKQNTLATIDADENQSLAPNPNYPSKXLAPIA---LNAQKFSPTITWYN 437
DB 1257 GQVTVGEAVTFSVID---NILSKI---ENEYEVLYLKLACGVRSLSKQLENNVMTFVN 1310
QY 438 -----QFLELEKTKQLRLDQV-YGNIAVNFENGVRVDTGSNWSEVL 481
DB 1311 VKDILNSRFKNKFNQV-LESDLIPYKDLTSSNVVVKDPYKFLNKEKRRDKFLSSYNIK 1369
QY 482 POIQETTARIIFNGKDLNVERRIAANVPSDPL-----ETTKPDM-TLKEALKIAPGF 533
DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSKSDLSIKYINDKQGE 1410
QY 534 NEPN-----GNLQYQKQKITE-----FDNFQOTSONIKNLQELNATNIYT 576
DB 1411 NEKLPFLNNIETLYKTVNDKIDLFIHLEAKVLNITYEK---SNVEVKIKELN--YLKT 1465

QY 587 MNLIRDKRFHYDRNNAIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEI 646
DB 1 MNLVRDP-YHYDNGNIVGVDDSVLKAYKQILNWSDDGVSLNLDDEVDNALSGYMLQI 59
QY 647 EDETE-----GLKEVINDRYDMLNLSLRQDGKTFIDPKKYNDKPLVYISNPN 693
DB 60 KKPNSHLTNSPVTTITLAGKDSGVGLYRVL-----DAGFLDPNKFENWRSVLV-DPG 112
QY 694 KYVNVYAVTKEN-TIINPSENGDTSTNGIKKILIPSKGYEI 734
DB 113 DDVYVYAVTKEDFNAVTRDENGNI-A-NKLKNTLVLSGKIKEI 153

RESULT 4
G59104
hypothesis: protein pXOI-111 - Bacillus anthracis virulence plasmid pXOI
C/Species: Bacillus anthracis
C/Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-May-2000
C/Accession: G59104
R/Okinaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler, J. Bacteriol. 181, 6509-6515, 1999
A/Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbored by B. anthracis strain Sterne
A/Reference number: A59091; MUID:99445483; PMID:10515943
A/Accession: G59104
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-204 <OKI>
A/Cross-references: GB:AF065404; NID:G4894216; PIDN:AA032415.1; PID:G4894327
A/Experimental source: strain Sterne
A/Note: similar to hypothetical protein in the protective antigen domain; ypa, plasmid pXOI
C/Genetics:
A/Gene: pXOI-111
A/Genome: plasmid

Query Match 6.2%; Score 235.5; DB 2; Length 204;
Best Local Similarity 34.6%; Pred. No. 2.6e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNLIRDKRFHYDRNNAIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEI 646
DB 1 MNLVRDP-YHYDNGNIVGVDDSVLKAYKQILNWSDDGVSLNLDDEVDNALSGYMLQI 59
QY 647 EDETE-----GLKEVINDRYDMLNLSLRQDGKTFIDPKKYNDKPLVYISNPN 693
DB 60 KKPNSHLTNSPVTTITLAGKDSGVGLYRVL-----DAGFLDPNKFENWRSVLV-DPG 112
QY 694 KYVNVYAVTKEN-TIINPSENGDTSTNGIKKILIPSKGYEI 734
DB 113 DDVYVYAVTKEDFNAVTRDENGNI-A-NKLKNTLVLSGKIKEI 153

RESULT 5
F82885
hypothesis: protein UUA482 [imported] - Ureaplasma urealyticum
C/Species: Ureaplasma urealyticum
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: F82885
R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caseall, G.H.
submitted to GenBank, February 2000
A/Description: The complete sequence of Ureaplasma urealyticum. Alternate views of a min
A/Reference number: A82870
A/Accession: F82885
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-468 <GLA>
A/Cross-references: GB:AE002145; GB:AF222894; NID:G68999476; PIDN:AAF30894.1; GSPDB:GN001
A/Experimental source: serovar 3; biovar 1
C/Genetics:
A/Gene: UUA482
A/Genetic code: SGC3

Query Match 5.4%; Score 202; DB 2; Length 4688;
Best Local Similarity 21.3%; Pred. No. 0.13;

Matches 174; Conservative 131; Mismatches 325; Indels 186; Gaps 40;
QY 19 LLGYFSDLNFPQPMVVTSTTGDLSIPSESELENIPSENQYFQSAIWSGFIKVKKSDYET 78
DB 3699 LVDVYLD-NIHQNDITRKIPKDHV-SKEIIEFPGVTMISKHGNWKSPTTTANFEFK 3756
QY 79 FATSADNHVTMVDDDEVINKASNNKIRLEKRIYQIKIQORENPTKGLDFKLYWTD 138
DB 3757 IETQ-----DDNDVLNNIDATVKFKDEHNNIKQIVRIKEN-----ND 3795
QY 139 SQNKKEVISSDNLQPLKQSSN-----SRKKSRSAGPTVDPD----- 179
DB 3796 WLKGGI-----DNLN-PETKYKLENTLSPLKTHNTLSINDKENISILITETGNPLKV 3951
QY 180 ---NDGIPDSLEVEGYTDVKNKRTFLSPFWSINIEHKKGLTKYKSSPEKWSPTASDYS 235
DB 3852 IOTONDINTOQTINVTLSGVNSK-YNGRKQKVYKDNVNNVIYESS---LITLQKQND 3907
QY 236 FEKVTRIGDKVNSPEARHPLVAAPYIVHVDMENILSKN-EDOSTONTSET-RTISKNT 293
DB 3908 YQLLSNLSN-----REYFEKIEIHNISNTNPFDEKLGUNGSFTITOKNT 3957
QY 294 STSRHTS-EVHGNAEVSHPF-----FDIGGSVSAGFS-----NSNS---STVAID 335
DB 3958 TVQNDSSATIVGTGWFNFKIKSEDKILENNQVVAWFAPKETITDNTMLQYTRPLK 4017
QY 336 HSLSLAGERTWATMGINT---ADTARLNANIRYNTGTAPIYXVLTSLVLGKQTLA 392
DB 4018 DVTSDPKEGFWADLSNSVNFKEETTKLVKIQFVNKPKAKNNINNSNNVLDNTSI 4077
QY 393 ---TIDADENQLSQILAPNNYPSKNLAPALNAQK-FSSTPITWY--NQFLELE 443
DB 4078 NSNYEFTKVGDKHLNITSSNNVNTSQTINFTLSGVKSVGKKILSKYKNDTSESI 4137
QY 444 KTKQLRLDTPQVGNATYFNGR-----VAVDTGSNWSEVLPOIQTETARIENGKD 497
DB 4138 HTNEVLIESNKTQYNTILLNNLKNRTYTLIDVKLIDNNVSDPFKEGNTNSFTITRTSA 4197
QY 498 LNLVERRIAAVNSDPLETT-----KPDMLTKEA-----LKIAFGFNEPENGLO 541
DB 4198 INVLTIEISNRSTNLKSTIIKINLNDPNVLRDQATIVYGNKNQKMGFTITVSGNIK 4257
QY 542 YQKDKTEFPDFPDQOTSONIK-NOLAEUNATNYTVLDKIKKAKONILIRDKRHYD- 599
DB 4258 YLTATLVDLNFN-DKVNIVNISFNKPSAAEN-----IGIDKSNII-----YNNDS 4304
QY 600 ---RNNIATVGA---DESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEIETE 650
DB 4305 IPKLEINNDIIVNGPINKELIVVKNANQK--NNIDVDLGLQINPKIAHLR-FLAKPKSTN 4361
QY 651 GLKEVIND--RYDMLNLSL-RQDGKTFIDF-----KKY----- 681
DB 4362 ---NDIETNVINGSLVNDGKTSIRFTLNNLKANKLYSLVDVYVLVNNNSNTIVE 4415
QY 682 NDKPLVYISNPNKYVAVTKENTINPSENGDTS 717
DB 4416 SNKLP-KLANNINYQIK---INKSHTII--SKNGEWS 4445

RESULT 6
E90598
membrane nuclease, lipoprotein [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C/Species: Mycoplasma pulmonis
C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C/Accession: E90598
R/Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Sanson, D.; Galisson, F.; Moszer, I.
Nucleic Acids Res. 29, 2145-2153, 2001
A/Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul
A/Reference number: A99512; MUID:21267165; PMID:11353084
A/Accession: E90598
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1125 <KUR>

QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVIMVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120
DB 90 QSAIWSGFIKVKKSDEYTFATSDADNHVIMVDDQEVINKASNSNKIRLEKGRLYQIKIQY 149
QY 121 QRENTEKGLDPKLYWTSQNKKEVSSDNQLPLKOKSSNSRKRKSTAGPTVPDRDN 180
DB 150 QRENTEKGLDPKLYWTSQNKKEVSSDNQLPLKOKSSNSRKRKSTAGPTVPDRDN 209
QY 181 DGIPOSLEVEGYTVQNKRTFLSPWISNIHEKGLTKYKSPKWSASDPYDFEKT 240
DB 210 DGIPOSLEVEGYTVQNKRTFLSPWISNIHEKGLTKYKSPKWSASDPYDFEKT 269
QY 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
DB 270 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 329
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
DB 330 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 389
QY 361 NANIRYVNTGTAPIYVNLPTTSLVIGKQTLATIDADENQSLIAPNNYPSKNLAPIA 420
DB 390 NANIRYVNTGTAPIYVNLPTTSLVIGKQTLATIDADENQSLIAPNNYPSKNLAPIA 449
QY 421 LNAOKKFSPTITMNYQNFLEKTKQLRLDQVYGNIAVNFENGVRVDTGNSKSEV 480
DB 450 LNAOKKFSPTITMNYQNFLEKTKQLRLDQVYGNIAVNFENGVRVDTGNSKSEV 509
QY 481 LPQIQETTARIIFNGKDLNVERIAAVNPSPLETTKPDMLKALKIAGFNEPNGL 540
DB 510 LPQIQETTARIIFNGKDLNVERIAAVNPSPLETTKPDMLKALKIAGFNEPNGL 569
QY 541 QYQKDIITFDNFDOOTQNTKQALAEALNATNYVLDKIKLNKAMILLIRKPFHYDR 600
DB 570 QYQKDIITFDNFDOOTQNTKQALAEALNATNYVLDKIKLNKAMILLIRKPFHYDR 629
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKRKILSGYIVIEDEGLKEVINDRY 660
DB 630 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKRKILSGYIVIEDEGLKEVINDRY 689
QY 661 DMLNISLRODGTFTDFKXNDKPLIYSNPNYKVNVAVTKNTIINPSENGSTNG 720
DB 690 DMLNISLRODGTFTDFKXNDKPLIYSNPNYKVNVAVTKNTIINPSENGSTNG 749
QY 721 IKKILIFSKGYBIG 735
DB 750 IKKILIFSKGYBIG 764

RESULT 2
I40862
iota toxin component Ib - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Oct-1999
C:Accession: I40862; S42774
R:Perelle, S.; Gibert, M.; Boquet, P.; Popoff, M.R.
Infect. Immun. 61, 5147-5156, 1993
A:Title: Characterization of Clostridium perfringens iota-toxin genes and expression in
A:Reference number: I40861; MUID:94041637; PMID:8225592
A:Accession: I40862
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-875 <RES>
A:Cross-references: EMBL:X73562; NID:G929031; PIDN:CAA51960.1; PID:G9414655

Query Match 23.1%; Score 871; DB 2; Length 875;
Best Local Similarity 31.1%; Pred. No. 1.5e-36;
Matches 252; Conservative 131; Mismatches 278; Indels 150; Gaps 31;

QY 1 EVQKQENPLNESSSSGLLYGYFSDNLPQAPMVVTSSTGTGLSIFSSLENIIPSE-NOY 59
DB 35 DTNQKEBITNLTSSNGLMGYFADEHFHFXDLEMAPIKNGDLKFEKKVYDKLTEDNSS 94

QY 60 FOSAIWSGFIKVKKSDEYTFATSDADNHVIMVDDQEVINKASNSNKIRLEKGRLYQIKIQY 119
DB 95 IKSIRWTGRIIPSEDEGYILSTDR-NDVLQINAKGDIKAK--TLKVNKKQAYNIRIE 150
QY 120 YQREN-----PTEKGLDFKLYWTSQNKKEVSSDNQLPLKOKSSNSRKRKST 170
DB 151 IQDNKLGSDNLSVP-----KLW-ELUNGKTVIPEENLFFRDYSKIDEND----- 195
QY 171 AGPTVP-----DRNDGIPDSLEVEGYTVQNKRTFLSPWISNIH 211
DB 196 --PFIENNNFFDVRFFSAAWEDDLTDNDNIPDAVEKNGYTI---KDSIAVKWNSFA 249
QY 212 EKKGLTKYKSSPEKWSASDPYDFEKTGRIDKNVSPKARHPLVAAYPIVHVDMENIIL 271
DB 250 E-QCYKKYVSSYLESTAGDPYDYQKASGSDKAIKLEARDPLVAAYPVGVGEMENLII 308
QY 272 SKNEDOSTQNTDSETRISKNSTSRHTSEVHGNAEVAHSPFDIGGSVAGFSNSNST 331
DB 309 STNEHASS---DQKTVSRAITNSKTANTV-----GVSSISAGYQNGFTCN 351
QY 332 VAIDHS-----LSLAGERTWAEWTGLNTADTARLNNIRYVNTGTAPIYVNLPTTSL 383
DB 352 ITTSYSHTTDNSTPAVDQSGESNMTGLSINKGESAYINANVRYNTGTAPMYKVTPTNL 411
QY 384 VLGNQTLATIDADENQSLIAPNNYPSKNLAPIALNAOKKFSPTITMNYQNFLELE 443
DB 412 VL-DGETLATIKADNQIGNNLSPNETYPKGLSPLALNTMDFNARLIPINYDQKKLD 470
QY 444 KTKQLRLDQVYGNIAVNFENGVRVDTGNSKSEVLPQIQETTARIIFNGKDLNVER 503
DB 471 SGQIKLETTQVSGNTGTYKN-SQGI-I-TEGNSWSNYISQIDSVSASIIID-TGSGTFR 527
QY 504 RIAAVNPSPLETTKPDMLKALKIAGFNEPNGLQY-QGKDIITE--FDNFDOOTSQ 560
DB 528 RVAARKQGNPEDT-PEITIGEAIKAFSATK-NGELLYFNGIPIDESCVELIFDDNTSE 585
QY 561 NIKNQALAEALNATNYVLDKIKLNKAMILLIRKRF--HYDR-NNIAVGADESVMKEAHR 617
DB 586 IIEQQLKYLDDKKIYVNV---KLERGMNIIKVPYSYFTNFDEYNNFP--ASWSNIDTKQ 639
QY 618 EVINSSTEGI-----LLNIDKDKRKILSGY-----IVEIEDTEGLKEVIND 658
DB 640 DGIQSVANKLSGTEKIIIPMSKLPKYRVFSGYKDPSTSNSTIWNKSKQKTDYLP 699
QY 659 RYDMLNIS-----SLRQDQGTFTDFKXND--KLPLYISNPNYKVNVAVTKNTIIN 696
DB 700 EKDYTKFSYEFETTKDSSDIEITLTSSGVIFLDNLSITELNSTPEILKEPKVPSDOB 759
QY 697 -----NVAVTKNTIINPSENGSTNGI 721
DB 760 ILDAHNNKYADIKLDI-----NTGNTYIDGI 785

RESULT 3
I39933
cryptic protein - Bacillus anthracis
C:Species: Bacillus anthracis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999
C:Accession: I39933
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr
A:Reference number: I39933; MUID:89172073; PMID:3148491
A:Accession: I39933
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-192 <RES>
A:Cross-references: GB:M22589; NID:G143280; PIDN:AAA22636.1; PID:G143281

Query Match 6.2%; Score 235.5; DB 2; Length 192;
Best Local Similarity 34.6%; Pred. No. 2.4e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 13.0855 Seconds
(without alignments)
5403.004 Million cell updates/sec

Title: US-09-848-909a-11

Perfect score: 3774
Sequence: 1 EVKQENLLNESBSSQGLL.....TSTNGIKLILFSKKGVEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3748	99.3	764	2 I39934	protective antigen
2	871	23.1	875	2 I40862	iota toxin compo
3	235.5	6.2	192	2 I39933	cryptic protein -
4	235.5	6.2	204	2 G59104	hypothetical prote
5	202	5.4	4688	2 F82885	hypothetical prote
6	191.5	5.1	1125	2 E92058	membrane nucleas
7	190.5	5.0	1639	2 S05603	major merozoite su
8	185.5	4.9	2529	2 B64635	toxin-like outer m
9	184.5	4.9	2401	2 T28676	rhothry protein -
10	182	4.8	6713	2 B89921	hypothetical prote
11	181	4.8	752	2 G90599	rhothry protein -
12	181	4.8	2269	2 T28677	hypothetical prote
13	179.5	4.8	1072	2 A86827	surface-located me
14	178	4.7	1302	1 JC6009	hypothetical prote
15	178	4.7	4919	2 T31105	Pil5 protein - Myc
16	177.5	4.7	979	2 JQ0894	uncharacterized pr
17	176	4.7	1227	2 C97033	major merozoite su
18	175	4.6	1631	1 SAZOK1	hemolysin [import
19	174	4.6	1635	2 A10452	toxin-like outer m
20	173.5	4.6	2399	2 H71879	botulinum neurotox
21	173	4.6	1193	2 S68218	hypothetical prote
22	173	4.6	1658	2 S55101	alpha-toxin - Clos
23	173	4.6	2178	2 S55805	hypothetical prote
24	170.5	4.5	786	2 T18469	hypothetical prote
25	170.5	4.5	821	2 S67087	SMC1 protein homol
26	170.5	4.5	1230	2 S56850	cell surface antig
27	170.5	4.5	2340	2 B71704	filamentous hemaag
28	170	4.5	4152	2 T31102	hypothetical prote
29	170	4.5	5005	2 F82884	hypothetical prote

30 169 4.5 1939 2 T18372 repeat organellar
31 168 4.5 3724 2 T18427 hypothetical prote
32 166.5 4.4 1115 2 T41342 probable coiled-co
33 166 4.4 1837 2 T41023 probable nuclear p
34 165 4.4 1516 2 E71619 RAD2 endonuclease
35 165 4.4 2013 2 AD1129 probable peptidogl
36 164.5 4.4 1365 2 T30822 cell wall-associat
37 164.5 4.4 2334 2 S32920 serine proteinase
38 164 4.3 769 2 F98870 ORF MSV156 hypothe
39 164 4.3 1127 2 T28317 hypothetical prote
40 164 4.3 1272 2 C90593 fibrinogen-binding
41 164 4.3 1315 2 T28679 surface membrane p
42 163.5 4.3 624 2 P66003 actin-interacting
43 163 4.3 1033 2 T37715 probable major sur
44 163 4.3 1640 2 A24594 DNA-directed RNA p
45 163 4.3 2910 2 T28156

ALIGNMENTS

RESULT 1

I39934
Protective antigen precursor - Bacillus anthracis plasmid
C/Species: Bacillus anthracis
C/Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 01-Dec-2000
C/Accession: I39934; S69160; F59104
R/Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A/Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr
A/Reference number: I39933; MUID:89172073; PMID:3148491
A/Accession: I39934
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-764 <RES>
A/Cross-references: GB:M22589; NID:g143280; PIDN:AAA22637.1; PID:g143282
R/Friedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klimpel, K.R.; Birch, N.P.; Loh, Y.P.
Arch. Biochem. Biophys. 316, 5-13, 1995
A/Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SF
A/Reference number: S69160; MUID:95142670; PMID:7840657
A/Accession: S69160
A/Molecule type: protein
A/Residues: 197-202 <FRI>
R/Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle
J. Bacteriol. 181, 6509-6515, 1999
A/Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbor
A/Reference number: A59091; MUID:99445483; PMID:10515943
A/Accession: F59104
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-313, 'Q', 315-764 <OKI>
A/Cross-references: GB:AF065404; NID:g4894216; PIDN:AAD32414.1; PID:g4894326
A/Experimental source: strain Sterne
A/Note: similar to anthrax toxin moiety, protective antigen, paga formerly pag, plasmid
C/Genetics:
A/Gene: pXO1-110
A/Genome: plasmid
C/Function:
A/Description: three component exotoxin; protective antigen binds to receptors on the s
y active components edema factor or lethal factor; the complex is internalized by recep
C/Keywords: exotoxin.
F/1-29/Domain: signal
F/30-196/Domain: propeptide #status predicted <PRO>
F/197-202/Product: protective antigen #status experimental <MAT>
Query Match 99.3%; Score 3748; DB 2; Length 764;
Best Local Similarity 99.5%; Pred. No. 2.3e+182;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENLLNESBSSQGLLYYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYQF 60
DB 30 EVKQENLLNESBSSQGLLYYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYQF 89

181 DGIIPDSLEVEGYVDVKNKRTFLSPWISNIEHKKGLTKYKSSPEKMWSTASDPYDFEVT 240
181 DGIIPDSLEVEGYVDVKNKRTFLSPWISNIEHKKGLTKYKSSPEKMWSTASDPYDFEVT 240
241 GRIDKNVSPPEARHPLVAAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
241 GRIDKNVSPPEARHPLVAAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLTADTARL 360
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLTADTARL 360
361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
421 LNAQKFSSTPTIMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVDGTSNWSEV 480
421 LNAQKFSSTPTIMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVDGTSNWSEV 480
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540
541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKONLILIRKRFHYDR 600
541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKONLILIRKRFHYDR 600
601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
661 DMLNISSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
661 DMLNISSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
721 IKKILIFSCKGYEIG 735
721 IKKILIFSCKGYEIG 735

Search completed: May 3, 2004, 19:57:42
Job time : 177.546 secs

241 GRIDKNVSPPEARHPLVAAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
241 GRIDKNVSPPEARHPLVAAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLTADTARL 360
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLTADTARL 360
361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
421 LNAQKFSSTPTIMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVDGTSNWSEV 480
421 LNAQKFSSTPTIMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVDGTSNWSEV 480
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540
541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKONLILIRKRFHYDR 600
541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKONLILIRKRFHYDR 600
601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
661 DMLNISSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
661 DMLNISSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
721 IKKILIFSCKGYEIG 735
721 IKKILIFSCKGYEIG 735

Query Match 99.3%; Score 3749; DB 1; Length 735;
Best Local Similarity 99.5%; Pred. No. 2.4e-296;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
1 EVKQENRLNSESSESSQGLGYVFDLNFQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
1 EVKQENRLNSESSESSQGLGYVFDLNFQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
61 QSAIWSGFIKVKSDVETATSDNHNVTMVDQEVINKASNSNKIRLKGRLYQIKIY 120
61 QSAIWSGFIKVKSDVETATSDNHNVTMVDQEVINKASNSNKIRLKGRLYQIKIY 120
121 QRENPTKGLDPKLYIWTDSQNKKEVSSDNLQJPELKQKSSNRKRSSTAGTVPDRDN 180
121 QRENPTKGLDPKLYIWTDSQNKKEVSSDNLQJPELKQKSSNRKRSSTAGTVPDRDN 180

RESULT 15
PCT-US03-35733-6
; Sequence 6, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-6

```

361 NANIRYVNTGAPINYNVLPTTSLVLGKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
361 NANIRYVNTGAPINYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
421 LNAQKFFSPTITMNYNOFLEKTKQLRLDQVYGNIAIYFNGRVRVDTGSNWSEV 480
421 LNAQKFFSPTITMNYNOFLEKTKQLRLDQVYGNIAIYFNGRVRVDTGSNWSEV 480
481 LPOIQTETARIIFNGKOLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 540
481 LPOIQTETARIIFNGKOLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 540
541 QYQCKDITEFDFNFDOQTSQNIKNQLAELNATNIYTVLDDKIKNAKNMLIRDKGFHYDR 600
541 QYQCKDITEFDFNFDOQTSQNIKNQLAELNATNIYTVLDDKIKNAKNMLIRDKGFHYDR 600
601 NNIAVGADESUVKEAHEVINSSTGLLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
601 NNIAVGADESUVKEAHEVINSSTGLLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNYYAVTKENTIIINPSENGDTSTNG 720
661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNYYAVTKENTIIINPSENGDTSTNG 720
721 IKKILIFSKKGYEIG 735
721 IKKILIFSKKGYEIG 735

RESULT 14
US-09-848-909A-7
; Sequence 7, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; PRIOR FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-7

Query Match 99.4%; Score 3750; DB 23; Length 735;
Best Local Similarity 99.5%; Pred. No. 2e-296;
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENOYF 60
DB 1 EVKQENRLNESSSSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMVDQEVINKASNSKNIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMVDQEVINKASNSKNIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKKGSTAGPTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKKGSTAGPTVPDRDN 180
QY 181 DGPIDSLVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWKSTADPSDFPEKVT 240
DB 181 DGPIDSLVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWKSTADPSDFPEKVT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNHEQSTONTSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNHEQSTONTSETRTISKNTSRTHT 300

361 NANIRYVNTGAPINYNVLPTTSLVLGKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
361 NANIRYVNTGAPINYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
421 LNAQKFFSPTITMNYNOFLEKTKQLRLDQVYGNIAIYFNGRVRVDTGSNWSEV 480
421 LNAQKFFSPTITMNYNOFLEKTKQLRLDQVYGNIAIYFNGRVRVDTGSNWSEV 480
481 LPOIQTETARIIFNGKOLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 540
481 LPOIQTETARIIFNGKOLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 540
541 QYQCKDITEFDFNFDOQTSQNIKNQLAELNATNIYTVLDDKIKNAKNMLIRDKGFHYDR 600
541 QYQCKDITEFDFNFDOQTSQNIKNQLAELNATNIYTVLDDKIKNAKNMLIRDKGFHYDR 600
601 NNIAVGADESUVKEAHEVINSSTGLLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
601 NNIAVGADESUVKEAHEVINSSTGLLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNYYAVTKENTIIINPSENGDTSTNG 720
661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNYYAVTKENTIIINPSENGDTSTNG 720
721 IKKILIFSKKGYEIG 735
721 IKKILIFSKKGYEIG 735

RESULT 13
US-09-848-909A-7
; Sequence 7, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; PRIOR FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-7

Query Match 99.4%; Score 3750; DB 23; Length 735;
Best Local Similarity 99.5%; Pred. No. 2e-296;
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENOYF 60
DB 1 EVKQENRLNESSSSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMVDQEVINKASNSKNIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMVDQEVINKASNSKNIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKKGSTAGPTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKKGSTAGPTVPDRDN 180
QY 181 DGPIDSLVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWKSTADPSDFPEKVT 240
DB 181 DGPIDSLVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWKSTADPSDFPEKVT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNHEQSTONTSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNHEQSTONTSETRTISKNTSRTHT 300
```

```
Db 481 LPOIETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGNEPENG 540
Qy 541 QYQKDIETDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAXNNILIRKRFHYDR 600
Db 541 QYQKDIETDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAXNNILIRKRFHYDR 600
Qy 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660
Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPYKVNVAVTYKNTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPYKVNVAVTYKNTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 11
US-09-848-909A-13
; Sequence 13, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-13

Query Match 99.4%; Score 3753; DB 23; Length 735;
Best Local Similarity 99.6%; Pred. No. 1.1e-296;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPDLKQKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPDLKQKSSNSRKRSTSGPTVPDRDN 180
Qy 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240
Db 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVIGKNOTLATIADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVIGKNOTLATIADENQLSQILAPNNYPSKNLAPIA 420
```

```
Qy 421 LNAQKKFSTPTITNYNQFLEKTKQLRLDQVYGNIAATYFENGVRVDTGNSHSEV 480
Db 421 LNAQKKFSTPTITNYNQFLEKTKQLRLDQVYGNIAATYFENGVRVDTGNSHSEV 480
Qy 481 LPOIETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGNEPENG 540
Db 481 LPOIETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGNEPENG 540
Qy 541 QYQKDIETDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAXNNILIRKRFHYDR 600
Db 541 QYQKDIETDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAXNNILIRKRFHYDR 600
Qy 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660
Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPYKVNVAVTYKNTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPYKVNVAVTYKNTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 12
PCT-US03-35733-7
; Sequence 7, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-7

Query Match 99.4%; Score 3750; DB 1; Length 735;
Best Local Similarity 99.5%; Pred. No. 2e-296;
Matches 731; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPDLKQKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPDLKQKSSNSRKRSTSGPTVPDRDN 180
Qy 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240
Db 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
```


Db 601 NNTAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISLRODQGTFFDFKYNKDLPLYISNPNYKVNVAVTENTIIINPSENGDTSTNG 720
Db 661 DMLNISLRODQGTFFDFKYNKDLPLYISNPNYKVNVAVTENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYSIG 735
Db 721 IKKILIFSKKGYSIG 735

RESULT 9
US-09-848-909A-8
; Sequence 8, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-8

Query Match 99.5%; Score 3754; DB 23; Length 735;
Best Local Similarity 99.6%; Pred. No. 9.5e-297;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSSQGLLGYFFSDFLNFOAPMVVTSSTTGDLSIPSSLENIPISENQYF 60
Db 1 EVKQENRLNESSSSQGLLGYFFSDFLNFOAPMVVTSSTTGDLSIPSSLENIPISENQYF 60
Qy 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNRKKRSTAGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNRKKRSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
Qy 241 GRIDKQVSPARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRTH 300
Db 241 GRIDKQVSPARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRTH 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLTADTARL 360
Qy 361 NANIRVNTGTAPIYVNLPTTSLVGLKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYVNLPTTSLVGLKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFFSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDVTGSNWSEV 480
Db 421 LNAQKFFSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDVTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENG 540
Qy 541 QYQKDIITDFDFNDQTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDRPHYDR 600

Db 541 QYQKDIITDFDFNDQTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDRPHYDR 600
Qy 601 NNTAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNTAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISLRODQGTFFDFKYNKDLPLYISNPNYKVNVAVTENTIIINPSENGDTSTNG 720
Db 661 DMLNISLRODQGTFFDFKYNKDLPLYISNPNYKVNVAVTENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYSIG 735
Db 721 IKKILIFSKKGYSIG 735

RESULT 10
PCT-US03-35733-13
; Sequence 13, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-13

Query Match 99.4%; Score 3753; DB 1; Length 735;
Best Local Similarity 99.6%; Pred. No. 1.1e-296;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSSQGLLGYFFSDFLNFOAPMVVTSSTTGDLSIPSSLENIPISENQYF 60
Db 1 EVKQENRLNESSSSQGLLGYFFSDFLNFOAPMVVTSSTTGDLSIPSSLENIPISENQYF 60
Qy 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNRKKRSTAGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNRKKRSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
Qy 241 GRIDKQVSPARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRTH 300
Db 241 GRIDKQVSPARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRTH 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLTADTARL 360
Qy 361 NANIRVNTGTAPIYVNLPTTSLVGLKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYVNLPTTSLVGLKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFFSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDVTGSNWSEV 480
Db 421 LNAQKFFSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDVTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENG 540


```
RESULT 7
US-09-848-909A-2
; Sequence 2, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-2

Query Match          99.5%; Score 3755; DB 23; Length 735;
Best Local Similarity 99.6%; Pred. No. 7.9e-297;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPEKGLDFKLYWTDSONKKEVITSSDNLOLPELKQSSNSRKRSTAGTVPDRDN 180
Db 121 QRENPEKGLDFKLYWTDSONKKEVITSSDNLOLPELKQSSNSRKRSTAGTVPDRDN 180
Qy 181 DGIPOSLEVEGYTDVYKNTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
Db 181 DGIPOSLEVEGYTDVYKNTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
Qy 241 GRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSTRISKNTSRTHT 300
Db 241 GRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSTRISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAGTGLNTADTARL 360
Qy 361 NANIRVNTGTAPIYNVLPPTSLVIGKQTLATIDADENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYNVLPPTSLVIGKQTLATIDADENQSLQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNIAATYNFENGRVVDTSNWSSEV 480
Db 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNIAATYNFENGRVVDTSNWSSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGNEPENG 540
Qy 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRPHYDR 600
Qy 601 NNAIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNAIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVTAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVTAVTKENTIINPSENGDTSTNG 720
```

```
RESULT 8
PCT-US03-35733-8
; Sequence 8, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-8
```

```
Query Match          99.5%; Score 3754; DB 1; Length 735;
Best Local Similarity 99.6%; Pred. No. 9.5e-297;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPEKGLDFKLYWTDSONKKEVITSSDNLOLPELKQSSNSRKRSTAGTVPDRDN 180
Db 121 QRENPEKGLDFKLYWTDSONKKEVITSSDNLOLPELKQSSNSRKRSTAGTVPDRDN 180
Qy 181 DGIPOSLEVEGYTDVYKNTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
Db 181 DGIPOSLEVEGYTDVYKNTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
Qy 241 GRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSTRISKNTSRTHT 300
Db 241 GRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSTRISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAGTGLNTADTARL 360
Qy 361 NANIRVNTGTAPIYNVLPPTSLVIGKQTLATIDADENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYNVLPPTSLVIGKQTLATIDADENQSLQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNIAATYNFENGRVVDTSNWSSEV 480
Db 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNIAATYNFENGRVVDTSNWSSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGNEPENG 540
Qy 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRPHYDR 600
Qy 601 NNAIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNAIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
```

```
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVTAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

; TITLE OF INVENTION: Modified Shine Dalgarno Sequences and Methods of Use Thereof
; FILE REFERENCE: PV595FCT
; CURRENT APPLICATION NUMBER: PCT/US03/19786
; CURRENT FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Mature PA sequence including an ETB signal sequence
PCT-US03-19786-4

Query Match 99.6%; Score 3758; DB 1; Length 735;
Best Local Similarity 99.6%; Pred. No. 4,5e-297;
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60

QY 61 QSAIWSGFTKVKKSDEYTFATSADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFTKVKKSDEYTFATSADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120

QY 121 QRENTEKGLDFKLWTDPSQNKKEVISSDNILQPELKQSSNRKSRSTAGTPVPDRN 180
DB 121 QRENTEKGLDFKLWTDPSQNKKEVISSDNILQPELKQSSNRKSRSTAGTPVPDRN 180

QY 181 DGIPDSLEVEGTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
DB 181 DGIPDSLEVEGTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

QY 301 SEVHGNAEVSASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEFTMGINTADTARL 360
DB 301 SEVHGNAEVSASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEFTMGINTADTARL 360

QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKFSSTPITMNYNQFLEKTKQLRLDQVYGNIAATYVNGRVRVDTGNSWSEV 480
DB 421 LNAQKFSSTPITMNYNQFLEKTKQLRLDQVYGNIAATYVNGRVRVDTGNSWSEV 480

QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIATGFNPNGL 540
DB 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIATGFNPNGL 540

QY 541 QYQKDIITEFDNFDOQTSQNKQLAELNATNIYTVLDKIKLNAKNMLIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSQNKQLAELNATNIYTVLDKIKLNAKNMLIRDKRPHYDR 600

QY 601 NNIAVGADESUVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
DB 601 NNIAVGADESUVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660

QY 661 DMLNSSLRQDGKTFIDFKKYNKPLPLYISNPNYKVNVAVTKENTIINPSENGDTSNG 720
DB 661 DMLNSSLRQDGKTFIDFKKYNKPLPLYISNPNYKVNVAVTKENTIINPSENGDTSNG 720

QY 721 IKKILIFSKGYEIG 735
DB 721 IKKILIFSKGYEIG 735

RESULT 6
PCT-US03-35733-2

; Sequence 2, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-2

Query Match 99.5%; Score 3755; DB 1; Length 735;
Best Local Similarity 99.6%; Pred. No. 7,9e-297;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60

QY 61 QSAIWSGFTKVKKSDEYTFATSADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFTKVKKSDEYTFATSADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120

QY 121 QRENTEKGLDFKLWTDPSQNKKEVISSDNILQPELKQSSNRKSRSTAGTPVPDRN 180
DB 121 QRENTEKGLDFKLWTDPSQNKKEVISSDNILQPELKQSSNRKSRSTAGTPVPDRN 180

QY 181 DGIPDSLEVEGTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
DB 181 DGIPDSLEVEGTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

QY 301 SEVHGNAEVSASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEFTMGINTADTARL 360
DB 301 SEVHGNAEVSASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEFTMGINTADTARL 360

QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKFSSTPITMNYNQFLEKTKQLRLDQVYGNIAATYVNGRVRVDTGNSWSEV 480
DB 421 LNAQKFSSTPITMNYNQFLEKTKQLRLDQVYGNIAATYVNGRVRVDTGNSWSEV 480

QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIATGFNPNGL 540
DB 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIATGFNPNGL 540

QY 541 QYQKDIITEFDNFDOQTSQNKQLAELNATNIYTVLDKIKLNAKNMLIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSQNKQLAELNATNIYTVLDKIKLNAKNMLIRDKRPHYDR 600

QY 601 NNIAVGADESUVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
DB 601 NNIAVGADESUVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660

QY 661 DMLNSSLRQDGKTFIDFKKYNKPLPLYISNPNYKVNVAVTKENTIINPSENGDTSNG 720
DB 661 DMLNSSLRQDGKTFIDFKKYNKPLPLYISNPNYKVNVAVTKENTIINPSENGDTSNG 720

QY 721 IKKILIFSKGYEIG 735
DB 721 IKKILIFSKGYEIG 735

```

; SEQ ID NO 10
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-10

Query Match          99.7%; Score 3761; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 2.5e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60
DB 1 EVKQENLLNESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDGYTFATSDAHVMTWDDQEVINKASNNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDGYTFATSDAHVMTWDDQEVINKASNNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNRKRGSTAGPTVPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNRKRGSTAGPTVPDRN 180
QY 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
DB 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYYPKSLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYYPKSLAPIA 420
QY 421 LNAQKFFSTPTIMYVNTQFLELEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
DB 421 LNAQKFFSTPTIMYVNTQFLELEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
QY 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
DB 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKKNTILIRDKREHYDR 600
DB 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKKNTILIRDKREHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVEIEDTEGLKEVINDRY 720
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVEIEDTEGLKEVINDRY 720
QY 721 IKKILIFSCKGYEIG 735
DB 721 IKKILIFSCKGYEIG 735
```

```

RESULT 4
US-09-848-909A-10
; Sequence 10, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
```

```

; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-10

Query Match          99.7%; Score 3761; DB 23; Length 735;
Best Local Similarity 99.7%; Pred. No. 2.5e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60
DB 1 EVKQENLLNESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDGYTFATSDAHVMTWDDQEVINKASNNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDGYTFATSDAHVMTWDDQEVINKASNNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNRKRGSTAGPTVPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNRKRGSTAGPTVPDRN 180
QY 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
DB 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYYPKSLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYYPKSLAPIA 420
QY 421 LNAQKFFSTPTIMYVNTQFLELEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
DB 421 LNAQKFFSTPTIMYVNTQFLELEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
QY 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
DB 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKKNTILIRDKREHYDR 600
DB 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKKNTILIRDKREHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVEIEDTEGLKEVINDRY 720
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVEIEDTEGLKEVINDRY 720
QY 661 DMLNSSLRQDGKTFIDFKKNDKPLIYSNPNYKVVAVTKENTIINPSENGDSTNG 720
DB 661 DMLNSSLRQDGKTFIDFKKNDKPLIYSNPNYKVVAVTKENTIINPSENGDSTNG 720
QY 721 IKKILIFSCKGYEIG 735
DB 721 IKKILIFSCKGYEIG 735
```

```

RESULT 5
PCT-US03-19786-4
; Sequence 4, Application PC/TUS0319786
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
```

```

US-09-848-909A-11
Query Watch      100.0%; Score 3774; DB 23; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.2e-298;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSELENIIPSENOYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSELENIIPSENOYF 60

61 QSAIWSGFIKVKKSDIEYFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDIEYFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

121 QRENTEKGLDPKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPDRN 180
DB 121 QRENTEKGLDPKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPDRN 180

181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTGYKSSPEKWSSTASDPYDFEKT 240
DB 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTGYKSSPEKWSSTASDPYDFEKT 240

241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420

421 LNAQKFSSTPIITMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480
DB 421 LNAQKFSSTPIITMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480

481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

541 QYQKDIITFDFNFDDQSTQNIKNQLAELNATYVTLVDKIKLNKAKNILLIRKRFHYDR 600
DB 541 QYQKDIITFDFNFDDQSTQNIKNQLAELNATYVTLVDKIKLNKAKNILLIRKRFHYDR 600

601 NNIAVGADESVMKEAREHREVINSTEGLLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAREHREVINSTEGLLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660

661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 3
US-09-848-909A-11
; Sequence 11, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis

```

```

US-09-848-909A-11
Query Watch      100.0%; Score 3774; DB 23; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.2e-298;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSELENIIPSENOYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSELENIIPSENOYF 60

61 QSAIWSGFIKVKKSDIEYFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDIEYFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

121 QRENTEKGLDPKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPDRN 180
DB 121 QRENTEKGLDPKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPDRN 180

181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTGYKSSPEKWSSTASDPYDFEKT 240
DB 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTGYKSSPEKWSSTASDPYDFEKT 240

241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420

421 LNAQKFSSTPIITMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480
DB 421 LNAQKFSSTPIITMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480

481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

541 QYQKDIITFDFNFDDQSTQNIKNQLAELNATYVTLVDKIKLNKAKNILLIRKRFHYDR 600
DB 541 QYQKDIITFDFNFDDQSTQNIKNQLAELNATYVTLVDKIKLNKAKNILLIRKRFHYDR 600

601 NNIAVGADESVMKEAREHREVINSTEGLLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAREHREVINSTEGLLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660

661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 3
PCT-US03-35733-10
; Sequence 10, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 175.546 Seconds
(without alignments)
4086.665 Million cell updates/sec

Title: US-09-848-909A-11

Perfect score: 3774

Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKILLFSSKKGHIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main: *

- 1: /cgn2_6/ptodata/2/paa/pctus COMB.pcp.*
- 2: /cgn2_6/ptodata/2/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/2/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/2/paa/US081 COMB.pcp.*
- 6: /cgn2_6/ptodata/2/paa/US082 COMB.pcp.*
- 7: /cgn2_6/ptodata/2/paa/US083 COMB.pcp.*
- 8: /cgn2_6/ptodata/2/paa/US084 COMB.pcp.*
- 9: /cgn2_6/ptodata/2/paa/US085 COMB.pcp.*
- 10: /cgn2_6/ptodata/2/paa/US086 COMB.pcp.*
- 11: /cgn2_6/ptodata/2/paa/US087 COMB.pcp.*
- 12: /cgn2_6/ptodata/2/paa/US088 COMB.pcp.*
- 13: /cgn2_6/ptodata/2/paa/US089 COMB.pcp.*
- 14: /cgn2_6/ptodata/2/paa/US090 COMB.pcp.*
- 15: /cgn2_6/ptodata/2/paa/US091 COMB.pcp.*
- 16: /cgn2_6/ptodata/2/paa/US092 COMB.pcp.*
- 17: /cgn2_6/ptodata/2/paa/US093 COMB.pcp.*
- 18: /cgn2_6/ptodata/2/paa/US094 COMB.pcp.*
- 19: /cgn2_6/ptodata/2/paa/US095 COMB.pcp.*
- 20: /cgn2_6/ptodata/2/paa/US096 COMB.pcp.*
- 21: /cgn2_6/ptodata/2/paa/US097A COMB.pcp.*
- 22: /cgn2_6/ptodata/2/paa/US097B COMB.pcp.*
- 23: /cgn2_6/ptodata/2/paa/US098 COMB.pcp.*
- 24: /cgn2_6/ptodata/2/paa/US099A COMB.pcp.*
- 25: /cgn2_6/ptodata/2/paa/US099B COMB.pcp.*
- 26: /cgn2_6/ptodata/2/paa/US100 COMB.pcp.*
- 27: /cgn2_6/ptodata/2/paa/US101 COMB.pcp.*
- 28: /cgn2_6/ptodata/2/paa/US102 COMB.pcp.*
- 29: /cgn2_6/ptodata/2/paa/US103 COMB.pcp.*
- 30: /cgn2_6/ptodata/2/paa/US104 COMB.pcp.*
- 31: /cgn2_6/ptodata/2/paa/US105 COMB.pcp.*
- 32: /cgn2_6/ptodata/2/paa/US106 COMB.pcp.*
- 33: /cgn2_6/ptodata/2/paa/US107 COMB.pcp.*
- 34: /cgn2_6/ptodata/2/paa/US108 COMB.pcp.*
- 35: /cgn2_6/ptodata/2/paa/US109 COMB.pcp.*
- 36: /cgn2_6/ptodata/2/paa/US110 COMB.pcp.*
- 37: /cgn2_6/ptodata/2/paa/US111 COMB.pcp.*
- 38: /cgn2_6/ptodata/2/paa/US112 COMB.pcp.*
- 39: /cgn2_6/ptodata/2/paa/US113 COMB.pcp.*
- 40: /cgn2_6/ptodata/2/paa/US114 COMB.pcp.*
- 41: /cgn2_6/ptodata/2/paa/US115 COMB.pcp.*
- 42: /cgn2_6/ptodata/2/paa/US116 COMB.pcp.*
- 43: /cgn2_6/ptodata/2/paa/US117 COMB.pcp.*
- 44: /cgn2_6/ptodata/2/paa/US118 COMB.pcp.*
- 45: /cgn2_6/ptodata/2/paa/US119 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1	3774	100.0	735	1	PCT-US03-35733-11	Sequence 11, Appl
2	3774	100.0	735	23	US-09-848-909A-11	Sequence 11, Appl
3	3761	99.7	735	1	PCT-US03-35733-10	Sequence 10, Appl
4	3761	99.7	735	23	US-09-848-909A-10	Sequence 10, Appl
5	3758	99.6	735	1	PCT-US03-19786-4	Sequence 4, Appl
6	3755	99.5	735	1	PCT-US03-35733-2	Sequence 2, Appl
7	3755	99.5	735	23	US-09-848-909A-2	Sequence 2, Appl
8	3754	99.5	735	1	PCT-US03-35733-8	Sequence 8, Appl
9	3754	99.5	735	23	US-09-848-909A-8	Sequence 8, Appl
10	3753	99.4	735	1	PCT-US03-35733-13	Sequence 13, Appl
11	3753	99.4	735	23	US-09-848-909A-13	Sequence 13, Appl
12	3750	99.4	735	1	PCT-US03-35733-7	Sequence 7, Appl
13	3750	99.4	735	23	US-09-848-909A-7	Sequence 7, Appl
14	3749	99.3	735	1	PCT-US03-35733-4	Sequence 4, Appl
15	3749	99.3	735	23	US-09-848-909A-4	Sequence 4, Appl
16	3749	99.3	735	23	US-09-848-909A-6	Sequence 6, Appl
17	3748	99.3	735	23	US-09-848-909A-5	Sequence 5, Appl
18	3748	99.3	735	1	PCT-US03-35733-5	Sequence 5, Appl
19	3748	99.3	735	1	PCT-US03-35733-19	Sequence 19, Appl
20	3748	99.3	735	1	PCT-US03-35733-20	Sequence 20, Appl
21	3748	99.3	735	1	PCT-US03-35733-21	Sequence 21, Appl
22	3748	99.3	735	22	US-09-791-537-43735	Sequence 43735, A
23	3748	99.3	735	23	US-09-848-909A-5	Sequence 5, Appl
24	3748	99.3	735	23	US-09-848-909A-19	Sequence 19, Appl
25	3748	99.3	735	23	US-09-848-909A-20	Sequence 20, Appl
26	3748	99.3	735	23	US-09-848-909A-21	Sequence 21, Appl
27	3748	99.3	735	30	US-10-410-647-30	Sequence 30, Appl
28	3748	99.3	736	1	PCT-US01-14372A-1	Sequence 1, Appl
29	3748	99.3	736	1	PCT-US01-14372A-2	Sequence 2, Appl
30	3748	99.3	736	1	PCT-US01-14372A-3	Sequence 3, Appl
31	3748	99.3	736	1	PCT-US01-14372A-4	Sequence 4, Appl
32	3748	99.3	736	1	PCT-US01-14372A-5	Sequence 5, Appl
33	3748	99.3	736	1	PCT-US01-14372A-6	Sequence 6, Appl
34	3748	99.3	736	1	PCT-US01-14372A-7	Sequence 7, Appl
35	3748	99.3	736	1	PCT-US01-14372A-8	Sequence 8, Appl
36	3748	99.3	736	1	PCT-US01-14372A-9	Sequence 9, Appl
37	3748	99.3	736	1	PCT-US01-14372A-10	Sequence 10, Appl
38	3748	99.3	736	1	PCT-US01-14372A-11	Sequence 11, Appl
39	3748	99.3	736	1	PCT-US01-14372A-12	Sequence 12, Appl
40	3748	99.3	736	1	PCT-US01-14372A-13	Sequence 13, Appl
41	3748	99.3	736	1	PCT-US01-14372A-14	Sequence 14, Appl
42	3748	99.3	736	1	PCT-US01-14372A-15	Sequence 15, Appl
43	3748	99.3	736	1	PCT-US01-14372A-16	Sequence 16, Appl
44	3748	99.3	736	1	PCT-US01-14372A-17	Sequence 17, Appl
45	3748	99.3	736	1	PCT-US01-14372A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

PCT-US03-35733-11
; Sequence 11, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-11

Query Match 100.0%; Score 3774; DB 1; Length 735;
Best Local Similarity 100.0%; Pred.No. 2.2e-298;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 1 EVQENRLLNESSSQGLGYYFSDLPQAPMVVTSSTTGDLSIPSSSELENIPEYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKRSTASGTPVDRDN 180
Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIDADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIKAKENQLSOILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFSSTPIITMNNYQFLEKTKQLRLDTPQVYGNATYNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKFSSTPIITMNNYQFLEKTKQLRLDTPQVYGNATYNFENGRVVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNPNGNL 540
Qy 541 QYQKDIETEPDFNPDQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPHYDR 600
Db 541 QYQKDIETEPDFNPDQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPHYDR 600
Qy 601 NNIAVGADESIVKEAAREVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Db 601 NNIAVGADESIVKEAAREVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

RESULT 15

```
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication No. US2002003958A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-14
```

```
Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 EVQENRLLNESSSQGLGYYFSDLPQAPMVVTSSTTGDLSIPSSSELENIPEYF 60
Db 1 EVQENRLLNESSSQGLGYYFSDLPQAPMVVTSSTTGDLSIPSSSELENIPEYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKRSTASGTPVDRDN 180
Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIDADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIKAKENQLSOILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFSSTPIITMNNYQFLEKTKQLRLDTPQVYGNATYNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKFSSTPIITMNNYQFLEKTKQLRLDTPQVYGNATYNFENGRVVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNPNGNL 540
Qy 541 QYQKDIETEPDFNPDQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPHYDR 600
Db 541 QYQKDIETEPDFNPDQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPHYDR 600
Qy 601 NNIAVGADESIVKEAAREVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Db 601 NNIAVGADESIVKEAAREVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

Search completed: May 3, 2004, 20:02:11
Job time : 38.6393 secs

121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKGSTASGPTVPDRD 180
121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKGSTASGPTVPDRD 180
181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIEHKKGLTKYKSSPEKWSASDPYDFEKT 240
181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIEHKKGLTKYKSSPEKWSASDPYDFEKT 240
241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
301 SEVHGNAEVHAFPDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
301 SEVHGNAEVHAFPDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
361 NANIRVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
361 NANIRVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
421 LNAQKFFSTPTIMYNOFLEKTKQLRLDQDVYGNIAATYFNPNRVRVDTGSNWSEV 480
421 LNAQKFFSTPTIMYNOFLEKTKQLRLDQDVYGNIAATYFNPNRVRVDTGSNWSEV 480
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540
541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKMNLIRDKRFHYDR 600
541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKMNLIRDKRFHYDR 600
601 NNTAVGADSVKAEHREVINSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
601 NNTAVGADSVKAEHREVINSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
721 IKKILPSKKGYEIG 735
721 IKKILPSKKGYEIG 735

RESULT 14
US-09-848-909-13
; Sequence 13, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-13

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60

181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIEHKKGLTKYKSSPEKWSASDPYDFEKT 240
241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
301 SEVHGNAEVHAFPDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
301 SEVHGNAEVHAFPDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
361 NANIRVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
361 NANIRVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
421 LNAQKFFSTPTIMYNOFLEKTKQLRLDQDVYGNIAATYFNPNRVRVDTGSNWSEV 480
421 LNAQKFFSTPTIMYNOFLEKTKQLRLDQDVYGNIAATYFNPNRVRVDTGSNWSEV 480
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540
541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKMNLIRDKRFHYDR 600
541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKMNLIRDKRFHYDR 600
601 NNTAVGADSVKAEHREVINSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
601 NNTAVGADSVKAEHREVINSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720
721 IKKILPSKKGYEIG 735
721 IKKILPSKKGYEIG 735

RESULT 13
US-09-848-909-12
; Sequence 12, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-12

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120

DB 361 NANIRYVNTGTAPIYNNVLPTTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAOKFSTPTITMNVNOFLEKTKQLRLDQVYGNATYVNGRVRVDTGSNWSEV 480
DB 421 LNAQDDFSTPTITMNVNGFLEKTKQLRLDQVYGNATYVNGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNYITVLDKIKLAKKNILIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNYITVLDKIKLAKKNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAGADESVVKEAHREVINSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735
RESULT 11
US-09-848-909-10
; Sequence 10, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-10
Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60
DB 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYEITFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYEITFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTSQNKKEVYSSDNQLPELKQSSNSRKRSTAGTPVPRDN 180
DB 121 QRENTEKGLDFKLYWTSQNKKEVYSSDNQLPELKQSSNSRKRSTAGTPVPRDN 180
QY 181 DGPISLVEGYTVQNKRTFLSPWISNIHEKGLTKYKSPEKWSASTASDPYDFEKT 240
DB 181 DGPISLVEGYTVQNKRTFLSPWISNIHEKGLTKYKSPEKWSASTASDPYDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNNVLPTTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYNNVLPTTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAOKFSTPTITMNVNOFLEKTKQLRLDQVYGNATYVNGRVRVDTGSNWSEV 480
DB 421 LNAQDDFSTPTITMNVNGFLEKTKQLRLDQVYGNATYVNGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNYITVLDKIKLAKKNILIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNYITVLDKIKLAKKNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAGADESVVKEAHREVINSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735
RESULT 12
US-09-848-909-11
; Sequence 11, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-11
Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60
DB 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYEITFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYEITFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTSQNKKEVYSSDNQLPELKQSSNSRKRSTAGTPVPRDN 180
DB 121 QRENTEKGLDFKLYWTSQNKKEVYSSDNQLPELKQSSNSRKRSTAGTPVPRDN 180
QY 181 DGPISLVEGYTVQNKRTFLSPWISNIHEKGLTKYKSPEKWSASTASDPYDFEKT 240
DB 181 DGPISLVEGYTVQNKRTFLSPWISNIHEKGLTKYKSPEKWSASTASDPYDFEKT 240

```
Db 541 QYQKDIETFDNFQOQTSQNIKNQALAEINATNIYTVLDKIKNAKKNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSGTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSGTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDQKTFIDFKYNDKPLVYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDQKTFIDFKYNDKPLVYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 9
US-09-848-909-8
; Sequence 8, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-8

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYTFATSDADNHVTMVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYTFATSDADNHVTMVDDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELQKSSNSRKKRSTSGPTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELQKSSNSRKKRSTSGPTVPDRN 180
QY 181 DGPDSLEVEGYVDVKNKKTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYVDVKNKKTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFIDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFIDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIDADENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIDADENQSLQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPIWNNYQFLEKTKQLRLDTQVYGNIAATYFNGRVRVDTGSNWSEV 480
Db 421 LNAQKFSSTPIWNNYQFLEKTKQLRLDTQVYGNIAATYFNGRVRVDTGSNWSEV 480
```

```
QY 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSPLETTKPDMTLKEALKIAFGNEPNGML 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSPLETTKPDMTLKEALKIAFGNEPNGML 540
QY 541 QYQKDIETFDNFQOQTSQNIKNQALAEINATNIYTVLDKIKNAKKNILIRDKRPHYDR 600
Db 541 QYQKDIETFDNFQOQTSQNIKNQALAEINATNIYTVLDKIKNAKKNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSGTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSGTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDQKTFIDFKYNDKPLVYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDQKTFIDFKYNDKPLVYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 10
US-09-848-909-9
; Sequence 9, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-9

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYTFATSDADNHVTMVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYTFATSDADNHVTMVDDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELQKSSNSRKKRSTSGPTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELQKSSNSRKKRSTSGPTVPDRN 180
QY 181 DGPDSLEVEGYVDVKNKKTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYVDVKNKKTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFIDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFIDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIDADENQSLQILAPNNYPSKNLAPIA 420
```

```
Db 721 IKKILIFSKGYEIG 735
|||||
RESULT 7
US-09-848-909-6
; Sequence 6, Application US/09848909
; Publication No. US2002039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-6

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDSYTATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSYTATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTASDPSEKVT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTASDPSEKVT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKULAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKULAPIA 420
Qy 421 LNAQKFFSSTPTIMYNNQFLELEKTKQLRLDTQVYGNATYNFENGRVRVDTGSNWSEV 480
Db 421 LNAQKFFSSTPTIMYNNQFLELEKTKQLRLDTQVYGNATYNFENGRVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Qy 541 QYQKDIITEFDNFDOQTQONIKQLAELNATNIYTVLDKIKLNKAKNILLIRKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTQONIKQLAELNATNIYTVLDKIKLNKAKNILLIRKRFHYDR 600
Qy 601 NNIAGADESVKAEHREINVESTGLLNDIKDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAGADESVKAEHREINVESTGLLNDIKDKIRKILSGYIVIEDTEGLKEVINDRY 660

Qy 661 DMLNISLRODGKTFIDFKYNDKPLIYISNPNYKVNVAVTKNTIINPSENGDSTNG 720
Db 661 DMLNISLRODGKTFIDFKYNDKPLIYISNPNYKVNVAVTKNTIINPSENGDSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 8
US-09-848-909-7
; Sequence 7, Application US/09848909
; Publication No. US2002039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-7

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDSYTATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSYTATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTASDPSEKVT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTASDPSEKVT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKULAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKULAPIA 420
Qy 421 LNAQKFFSSTPTIMYNNQFLELEKTKQLRLDTQVYGNATYNFENGRVRVDTGSNWSEV 480
Db 421 LNAQKFFSSTPTIMYNNQFLELEKTKQLRLDTQVYGNATYNFENGRVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Qy 541 QYQKDIITEFDNFDOQTQONIKQLAELNATNIYTVLDKIKLNKAKNILLIRKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTQONIKQLAELNATNIYTVLDKIKLNKAKNILLIRKRFHYDR 600
Qy 601 NNIAGADESVKAEHREINVESTGLLNDIKDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAGADESVKAEHREINVESTGLLNDIKDKIRKILSGYIVIEDTEGLKEVINDRY 660
```

; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-4

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDSYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDSYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 180
QY 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASDPYSDPEKVT 240
DB 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASDPYSDPEKVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTSKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTSKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSOILAPNNYPSKNLAPIA 420
DB 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSOILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480
DB 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLTKEALKIAGFNEPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLTKEALKIAGFNEPNGNL 540
QY 541 QYQKDIPTDFNPDQOSTQNIKNQLAELNATNIYVLDKIKLNAXNNILIRDKRPFHYDR 600
DB 541 QYQKDIPTDFNPDQOSTQNIKNQLAELNATNIYVLDKIKLNAXNNILIRDKRPFHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIPSKGYEIG 735
DB 721 IKKILIPSKGYEIG 735

RESULT 6

US-09-848-909-5

; Sequence 5, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-5

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDSYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDSYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 180
QY 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASDPYSDPEKVT 240
DB 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASDPYSDPEKVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTSKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTSKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSOILAPNNYPSKNLAPIA 420
DB 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSOILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480
DB 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLTKEALKIAGFNEPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLTKEALKIAGFNEPNGNL 540
QY 541 QYQKDIPTDFNPDQOSTQNIKNQLAELNATNIYVLDKIKLNAXNNILIRDKRPFHYDR 600
DB 541 QYQKDIPTDFNPDQOSTQNIKNQLAELNATNIYVLDKIKLNAXNNILIRDKRPFHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIPSKGYEIG 735

US-09-848-909-2

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5,1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRSTASGPTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRSTASGPTVPDRDN 180
QY 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKT 240
DB 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTIMYNOFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480
DB 421 LNAQKFSSTPTIMYNOFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPFNEPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPFNEPNGNL 540
QY 541 QYQKDIITEFDFNPDQQTSONIKQLAELNATNIYTVLDKIKLAKNKNILIRDKRPHYDR 600
DB 541 QYQKDIITEFDFNPDQQTSONIKQLAELNATNIYTVLDKIKLAKNKNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
DB 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPYKNVYAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPYKNVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 4

US-09-848-909-3

; Sequence 3, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-3

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5,1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRSTASGPTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRSTASGPTVPDRDN 180
QY 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKT 240
DB 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTIMYNOFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480
DB 421 LNAQKFSSTPTIMYNOFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPFNEPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPFNEPNGNL 540
QY 541 QYQKDIITEFDFNPDQQTSONIKQLAELNATNIYTVLDKIKLAKNKNILIRDKRPHYDR 600
DB 541 QYQKDIITEFDFNPDQQTSONIKQLAELNATNIYTVLDKIKLAKNKNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
DB 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPYKNVYAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPYKNVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 5

US-09-848-909-4

; Sequence 4, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment

```
Db 61 QSAIWGFIKVKKSDYTTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWTASDPYSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPARPLVAAPYPIVHVDMENILSKNEDQSTONTSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARPLVAAPYPIVHVDMENILSKNEDQSTONTSETRTISKNTSRTHT 300
QY 301 SEVHGNAEYHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEYHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIETDFDFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRFHYDR 600
Db 541 QYQKDIETDFDFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRFHYDR 600
QY 601 NNTAVGADESVMKEAREHREINSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNTAVGADESVMKEAREHREINSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRODGTFFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODGTFFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

RESULT 2

```
US-09-848-909-1
; Sequence 1, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 736
; TYPE: PRN
; ORGANISM: Bacillus anthracis
US-09-848-909-1
```

```
Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 EVAQENRLLNESSSQGLLYFFSDLNFPQAPVVTSSITGDLSPSSSELENIIPSENOYF 60
Db 1 EVAQENRLLNESSSQGLLYFFSDLNFPQAPVVTSSITGDLSPSSSELENIIPSENOYF 60
QY 61 QSAIWGFIKVKKSDYTTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Db 61 QSAIWGFIKVKKSDYTTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWTASDPYSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPARPLVAAPYPIVHVDMENILSKNEDQSTONTSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARPLVAAPYPIVHVDMENILSKNEDQSTONTSETRTISKNTSRTHT 300
QY 301 SEVHGNAEYHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEYHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIETDFDFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRFHYDR 600
Db 541 QYQKDIETDFDFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRFHYDR 600
QY 601 NNTAVGADESVMKEAREHREINSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNTAVGADESVMKEAREHREINSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRODGTFFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODGTFFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

RESULT 3

```
US-09-848-909-2
; Sequence 2, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 736
; TYPE: PRN
; ORGANISM: Bacillus anthracis
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:40:13 ; Search time 36.6393 Seconds
(without alignments)
5560.545 Million cell updates/sec

Title: US-09-848-909A-11

Perfect score: 3774

Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKKILFSSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3748	99.3	735	15	US-10-410-647-30
2	3748	99.3	736	12	US-09-848-909-1
3	3748	99.3	736	12	US-09-848-909-2
4	3748	99.3	736	12	US-09-848-909-3
5	3748	99.3	736	12	US-09-848-909-4
6	3748	99.3	736	12	US-09-848-909-5
7	3748	99.3	736	12	US-09-848-909-6
8	3748	99.3	736	12	US-09-848-909-7
9	3748	99.3	736	12	US-09-848-909-8
10	3748	99.3	736	12	US-09-848-909-9
11	3748	99.3	736	12	US-09-848-909-10
12	3748	99.3	736	12	US-09-848-909-11
13	3748	99.3	736	12	US-09-848-909-12
14	3748	99.3	736	12	US-09-848-909-13
15	3748	99.3	736	12	US-09-848-909-14

16	3748	99.3	736	12	US-09-848-909-15	Sequence 15, Appl
17	3748	99.3	736	12	US-09-848-909-16	Sequence 16, Appl
18	3748	99.3	736	12	US-09-848-909-17	Sequence 17, Appl
19	3748	99.3	736	12	US-09-848-909-18	Sequence 18, Appl
20	3748	99.3	736	12	US-09-848-909-19	Sequence 19, Appl
21	3748	99.3	736	12	US-09-848-909-20	Sequence 20, Appl
22	3748	99.3	736	12	US-09-848-909-21	Sequence 21, Appl
23	3748	99.3	736	12	US-09-848-909-22	Sequence 22, Appl
24	3748	99.3	736	12	US-10-442-502-7	Sequence 7, Appl
25	3748	99.3	763	15	US-10-442-502-5	Sequence 5, Appl
26	3748	99.3	764	15	US-10-442-502-6	Sequence 6, Appl
27	3745	99.2	735	12	US-10-402-466A-9	Sequence 9, Appl
28	3745	99.2	735	12	US-10-402-466A-13	Sequence 13, Appl
29	3745	99.2	735	14	US-10-332-282-13	Sequence 13, Appl
30	3741	99.1	735	12	US-09-848-909-30	Sequence 30, Appl
31	3741	99.1	764	12	US-10-253-286-681	Sequence 681, App
32	3741	99.1	764	9	US-10-245-871-681	Sequence 4, Appl
33	3641	96.5	764	13	US-10-106-014-4	Sequence 4, Appl
34	3641	96.5	764	13	US-10-106-695-4	Sequence 4, Appl
35	3641	96.5	764	14	US-10-105-694-4	Sequence 4, Appl
36	3037	80.5	599	12	US-09-848-909-24	Sequence 24, Appl
37	3032	80.3	595	14	US-10-332-282-11	Sequence 11, Appl
38	3032	80.3	573	12	US-10-402-466A-22	Sequence 22, Appl
39	2905	77.0	573	12	US-10-442-502-8	Sequence 8, Appl
40	2893	76.7	569	15	US-10-332-282-5	Sequence 5, Appl
41	2483	65.8	487	14	US-10-332-282-9	Sequence 9, Appl
42	2168	57.4	426	14	US-10-402-466A-24	Sequence 24, Appl
43	2122	56.2	423	12	US-10-332-282-7	Sequence 7, Appl
44	1619	42.9	318	14	US-10-332-282-3	Sequence 3, Appl
45	1338	35.5	258	14	US-10-332-282-7	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-410-647-30
; Sequence 30, Application US/10410647
; Publication No. US20030235818A1
; GENERAL INFORMATION:
; APPLICANT: PLEXUS VACCINE, INC.
; APPLICANT: Katritch, Vsevolod
; APPLICANT: Bordner, Andrew
; APPLICANT: Deans, Robert
; APPLICANT: Sumner, Mary
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES, AND METHOD OF IDENTIFYING SAME
; FILE REFERENCE: PLEX1110-1
; CURRENT APPLICATION NUMBER: US/10/410,647
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/373,668
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/371,256
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/371,250
; PRIOR FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-410-647-30

Query Match	99.3%	Score 3748	DB 15	Length 735
Best Local Similarity	99.5%	Pred. No. 5.1e-273		
Mismatches	731	Conservative	0	Gaps 0
Qy	1	EVKQENLLNESSSQGLLYFSDINFOAPMVVTSSTTGDLSIPSELENIPSENOYF	60	
Db	1	EVKQENLLNESSSQGLLYFSDINFOAPMVVTSSTTGDLSIPSELENIPSENOYF	60	
Qy	61	QSAWTSFGFKVKKDEYTFATSDNHNVTMWVDDQEVINKASNGKIRLEGRLYQIKIY	120	


```

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 884 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-471-044-5

Query Match      20.3%; Score 766; DB 2; Length 884;
Best Local Similarity 30.1%; Pred. No. 1.5e-46;
Matches 245; Conservative 129; Mismatches 295; Indels 146; Gaps 34;

QY 4 QENRLNSESQGLGYFSDLNFOAPMVVTSITGDSIPSELEN--IPSENYFQ 61
Db 42 QXNQ---QXEMRKGLGYFKGDF-SNLTFATRDSTLIYDQOTANKLLDKQEQYQ 97
QY 62 SAIWGFIVKKSDEYTFATSDNNHVTWVDDQEVINKASNSNKRIRLEKGRLYQIKYQ 121
Db 98 SIRWIGLIQSKETGFTNLSDEQAIIIEINGKIIISNKGKQVYVHLEKGLVPIKIEYQ 157
QY 122 RENPTEKGLD-----FKLYMTDSQNKKEVISSDNLQLPELKOKSS-----N 162
Db 158 SD--TKFNIDSKTFKELKLFKIDSONQPOQVQODELRNPEFNKESQEFLEKPSKINLFT 215
QY 163 SRKKSTSGAGTVPRDNDGIPDSLEVEGYTVDVNKKTFPLSPWISNHEKKGKGLTKYKS 222
Db 216 QXMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVKDDSL-ASKGYTKFVSN 267
QY 223 PEKWTASDPYDFEKVTVGRIDKNVSPPEARHPLVAAPIVHYVDMENIILSKNEDOSTQNT 282
Db 268 PLESHTVGDPYTDYEKAARDLDSNAKETFNPLVAFPSVNVSMKVLSPNENLS----- 323
QY 283 DSETRTISKNTSTRTHTSEVHNAFVHASFFDIGGSVSAGFSNSNS---TVADHSL 339
Db 324 ---NSVESHSSTNWSYT-----NTE-----CASVEAGIGPKGISFGVSVMYQHS 366
QY 340 LAGERTWAEWNG---LNTADIPARLANRYVNTGTAPIYVLPITTSVLGKNOTLATID 395
Db 367 VAQE--WGTSTGNTSFNTASAGYLNANRYNNVGTGALYDVKPTTSFVL-NDTIAIT 423
QY 396 ADENQLSQILAPNNYPSKNLAPIALNAQKFSSTPTMNYNQFLEKTKQLRLDTQV 455
Db 424 AKSNSTALNISPESYKKGQNGIAITSMDDFNHPITLNNKQVDNLLNNKPMLETNQT 483
QY 456 YGNIAVNFENGVRVDVTSNWEVLPOIQTETARIIFNGKDLNLVERIAAVNPSPDPLE 515
Db 484 DG---VYKIKDTRHNVITGGEWNGVIOQIKAKTASIIIVDDGE-RVAEKVAAKQYENPED 539
QY 516 TTKPDMTLKALKIAP--GFNEPENGLOQCKDITEFDF--NPDQOTSONIKNQLAEL-- 569
Db 540 KT-PSLTKDALKLSYDPEIKEIEGLLYKKNKEIYESSVMYLDENTAKEVTKQLNDTTG 598
QY 570 ---NATNIYTVLDKIKLNARONILIRDRPHYDRNNIAGADESVVKEAHRVINSSTEG 626
Db 599 KFKDVSHLYDV---KLTPOAVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGNG 650
QY 627 -----LLNLD-----KDIRKILSGYIVEIDTE-----GLKE 654
Db 651 KKQYSSNNPDPANLTNTDAQELKNKRDYVYISLYMKSEKNTQCEITIDGEIYPIITKYN 710
QY 655 VINDRYDMLN-----ISSLRQGGKTFIDFKYNDKPLIYISNPNYKVNVAATKE 704
Db 711 VNKNYKRLDIIAHNKSNPISLH-----IKTNDBITLFWDDDISI-TDVASIKPE 760
QY 705 NTIINPSENGDT-STNGIK---KILIFSKGYEIG 735
Db 761 N--LTDSEIKQIYRYGKLEKEDGILIDKGGIHYG 793

```

Search completed: May 3, 2004, 19:42:45
Job time : 18.0986 secs


```
QY 456 YGNIATYFNGRVRVDTGNSWSEVLPQIQTETARIIFNGKDLNVERIAAVNPSDPLE 515
Db 481 DG---VYKIKDTHGNIVTGGWNGVIOQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPED 536
QY 516 TTKPDMTLKEALKIAP--GFNEPNGNLQYQKDIETEDF--NPDQOTSONIKNQLAEL-- 569
Db 537 KT-PSLTLDKALKSYDPEIKIEBGLLYKKNKPIYESSVMTYLDENTAKEVTKQLNDITG 595
QY 570 ---NATNIYTVLKIKNAMNILLRDRPHYDRNNIAVCADESUVKEAHREVINSSTEG 626
Db 596 KFXDVSHLYDV---KLTPKMNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGG 647
QY 627 -----LLLNID-----KOIRKILSGYIVETEDTE-----GLKE 654
Db 648 KKOYSSNNPDANLTLNTDAQKLNKRDYIISLYMKSEKNTQCEITIDGIEYPIITKTVN 707
QY 655 VINDRYDMLNI--SSLRQDKTFIDFKYNDKPLIYSNPNYKVNKYAVTKENTINPSE 712
Db 708 VNKDNYKRLDIIAHNIKSNPISIIHKT-NDBITLFWDDISI-TDVASIKPEN--LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735
Db 764 IKQIYRYGKLEDDGILIDKKGIIHYG 790

RESULT 12
US-09-307-106-8
; Sequence 8, Application US/09307106
; Patent No. 6603063
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schweits, James
; APPLICANT: Loewer, David
; APPLICANT: Duilum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacey
; TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESS: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32608-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/307,106
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/073,898
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
```

```
; REFERENCE/DOCKET NUMBER: MA-708C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: PS177C8a
; US-09-307-106-8

Query Match 20.3%; Score 768; DB 4; Length 881;
Best Local Similarity 30.1%; Pred. No. 1,1e+46;
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;

QY 4 QENRLNNESSSSQGLLYGYSFNLNFOAPMVVTSSTGDLSPSSSELEN--IPSENQVPO 61
Db 39 QKIQ---QKEMDRKGLLYGFKGDF-SNLTWFAFTRDSTLYDOOTANKLLDKKQOEYQ 94
QY 62 SAIWSGFIKVKSDDEVTFATSDNHNVTMVDDQEVINKASNNKIRLEKGRLYQIKIQYQ 121
Db 95 SIRWIGLIQSKETGDFTNLSEDEQAIIEINGKIISNKGKEQVHVLEKGLVPIKIIYQ 154
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNLQIPELKQKSS-----N 162
Db 155 SD--TKFNIDSKTFKELKLFKIDSONQPOQOQDELNPEFNKESQEBFLAKPXSINLFT 212
QY 163 SRKRSTSGAPTVDRDNDGIPDSLEVEGYVDVKNKRTFLSPWISNTHKGLTKYKSS 222
Db 213 QMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264
QY 223 PEKWTASDPYDFEKVTKRIDKNVSPERHPLVAAPIVHVDMENIILSKNEDOSTQNT 282
Db 265 PLESHTVGDPTDYKAAARDLDSNAKETFNPLVAAFPSVNVSMKVLSPNENLS---- 320
QY 283 DSETRITISKNTSRTHTSEVHGNARVHASFPDIGSVSAGFSNSNS---TVAIDHSLS 339
Db 321 ----NSVESHSSTNWSYT-----NTF-----GASVEAGIGPKGISFGSVNYQHSFT 363
QY 340 LAGERWAEFMG----LNTADTARLANRINYNTGTAPIYVLTPTSLVLGKKNQTLATID 395
Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANVRVNVGTGAIYDVKPTTSFVL--NNDTIATIT 420
QY 396 ADENQLSQILAPNNYPSKNLAPIALNAOKFSSPTITMNYNQFLELEKTKOLRLDTQV 455
Db 421 AKSNSTALNISPGESYPKKGQNGIATSMDDDFNSHPITLKKQVDNLLNKKPMLETNQT 480
QY 456 YGNIATYFNGRVRVDTGNSWSEVLPQIQTETARIIFNGKDLNVERIAAVNPSDPLE 515
Db 481 DG---VYKIKDTHGNIVTGGWNGVIOQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPED 536
QY 516 TTKPDMTLKEALKIAP--GFNEPNGNLQYQKDIETEDF--NPDQOTSONIKNQLAEL-- 569
Db 537 KT-PSLTLDKALKSYDPEIKIEBGLLYKKNKPIYESSVMTYLDENTAKEVTKQLNDITG 595
QY 570 ---NATNIYTVLKIKNAMNILLRDRPHYDRNNIAVCADESUVKEAHREVINSSTEG 626
Db 596 KFXDVSHLYDV---KLTPKMNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGG 647
QY 627 -----LLLNID-----KOIRKILSGYIVETEDTE-----GLKE 654
Db 648 KKOYSSNNPDANLTLNTDAQKLNKRDYIISLYMKSEKNTQCEITIDGIEYPIITKTVN 707
QY 655 VINDRYDMLNI--SSLRQDKTFIDFKYNDKPLIYSNPNYKVNKYAVTKENTINPSE 712
Db 708 VNKDNYKRLDIIAHNIKSNPISIIHKT-NDBITLFWDDISI-TDVASIKPEN--LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735
```

Db 95 SIRWIGLQSKETGDFNFSLSEDEQAIIEINGKIIISNKGKQVHLEKGLVPIKIEYQ 154
Qy 122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNTLQPLKOKSS-----N 162
Db 155 SD--TKFNIDSKTFKELKFKIDSQNPQVQOQDELNPFENFKESQEFLLAKPSKINFT 212
Qy 163 SRKRSTASDPYSDPEKVTGRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDOSTQNT 282
Db 213 QMKREIDED---TDTGDSIPDLWEENGVTI---QNRIVAKWDDSL-ASKGYTKFVSN 264
Qy 223 PEKMWSTASDPYSDPEKVTGRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDOSTQNT 282
Db 265 PLESHTVGDPYTDYKAAARDLDSNAKETFPNLAAPFVSNVSKVILSPENLS---- 320
Qy 283 DSETRTISKNTSTRTSTHTEVHGNAEVHASFPIIDGVSAGFSNSNS---TVAIDHSL 339
Db 321 ---NSVSHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVMYCHSET 363
Qy 340 LAGERTWAETMG---LNTADTARLANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATID 395
Db 364 VAQE--WGTSTGNTSQNTASAGYLNANVRNNVGTGAIYDVKPTTSFVL--NNDTIAIT 420
Qy 396 ADENQLSQILAPNNYPSKNAPIALNAOKKFSSTPTMKYNOFLELEXTKQLRLDTPQV 455
Db 421 AKSNSTALNISPGESEYKPKGQNGIAITSMDDFNSHPITLKKQVDNLLNKPMLTETQ 480
Qy 456 YGNATATNFENGRVDTGNSWSEVLPOIQTETARIFNGKDLNLYVERIAAVNPSDPLE 515
Db 481 DG----VYKIKTHGNITVGWNGVIOIIRAKTASIIVDGGE--RVAEKVAAKDVENPED 536
Qy 516 TTKPDMTLKEALKIAF--GFNEPNNGNLQYQKQDITBPDF--NPDQOTSONIKNQLAEL-- 569
Db 537 KT-PSLTKDALKSYPDEIKETEGLLYKKNKPIYESSVMYTDENTAKEVTKQLNDITG 595
Qy 570 ---NATNIYTVLDKIKNAXMIIILDRKPHYDENNIACADESVVKEAHREVINSSREG 626
Db 596 KFKOVSHLDY---KUTPRKNTVIR--LSILYN---AESNDNSIGKWTNIVSGGNG 647
Qy 627 -----LLNLID-----KDIRKILGYVIEIDTE-----GLKE 654
Db 648 KKOYSSNNPDANLTLNDAOKLNKRDYIISLYMKSEKNTQCEITIDGIEIPIITKTWN 707
Qy 655 VINDRYDMNT--SSLRQDGKTEIDFKYNDKPLVLSNPKYVNYVATKENTLNPSE 712
Db 708 VKNYKXELDIANNIKSNPISSHIKT--NDEITLFWDDISI--TDVASIKPEN--ITDSE 763
Qy 713 NGDT-STNGIK---KILIFKKGVEIG 735
Db 764 IKQIYSRYGIKLEDGILIDKKGGIHYG 790

RESULT 11

US-09-073-898-32
; Sequence 32, Application US/09073898
; Patent No. 6242869

GENERAL INFORMATION:

; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schneff, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacey
; TITLE OF INVENTION: No. 6242869el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,898
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-708C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: PS177C8
; US-09-073-898-32

Query Match 20.3%; Score 768; DB 3; Length 881;

Best Local Similarity 30.1%; Pred. No. 1.1e-46;

Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;

Qy 4 QENRLNSESSESGQLGYFSDLNFPQAPMVYTSSTGDLSPSELEN--IPSENQYFQ 61
Db 39 QKQK---QKEMRKGLLYGFFKGF-SNLNFWATRDSTLLYDQOTANKLLDKKQEQY 94
Qy 62 SAIWGFIKVKSEDEYTFATSDNHNVTWDDQEVINKASNSKIRLEKRLYQIKQYQ 121
Db 95 SIRWIGLQSKETGDFNFSLSEDEQAIIEINGKIIISNKGKQVHLEKGLVPIKIEYQ 154
Qy 122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNTLQPLKOKSS-----N 162
Db 155 SD--TKFNIDSKTFKELKFKIDSQNPQVQOQDELNPFENFKESQEFLLAKPSKINFT 212
Qy 163 SRKRSTASDPYSDPEKVTGRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDOSTQNT 282
Db 213 QMKREIDED---TDTGDSIPDLWEENGVTI---QNRIVAKWDDSL-ASKGYTKFVSN 264
Qy 223 PEKMWSTASDPYSDPEKVTGRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDOSTQNT 282
Db 265 PLESHTVGDPYTDYKAAARDLDSNAKETFPNLAAPFVSNVSKVILSPENLS---- 320
Qy 283 DSETRTISKNTSTRTSTHTEVHGNAEVHASFPIIDGVSAGFSNSNS---TVAIDHSL 339
Db 321 ---NSVSHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVMYCHSET 363
Qy 340 LAGERTWAETMG---LNTADTARLANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATID 395
Db 364 VAQE--WGTSTGNTSQNTASAGYLNANVRNNVGTGAIYDVKPTTSFVL--NNDTIAIT 420
Qy 396 ADENQLSQILAPNNYPSKNAPIALNAOKKFSSTPTMKYNOFLELEXTKQLRLDTPQV 455
Db 421 AKSNSTALNISPGESEYKPKGQNGIAITSMDDFNSHPITLKKQVDNLLNKPMLTETQ 480

QY 237 EKVTGEIDKNVSPARHPPLVAAPIVHVDMENIILSKNEDOSTONTSETRTISKNTSTS 296
DB 241 EKVTGEIDKNVSPARHPPLVAAPIVHVDMENIILSKNEDOSTONTSETRTISKNTSTS 300
QY 297 RHTTSEVHGNAEVAHSAFFDGGSVAGFSNSNSTVAIDHSLSLAGERTWAEITMGLNTAD 356
DB 301 RHTTSEVHGNAEVAHSAFFDGGSVAGFSNSNSTVAIDHSLSLAGERTWAEITMGLNTAD 360
QY 357 TARLNNIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENOLSOILAPNYYPSKNL 416
DB 361 TARLNNIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQJSQLAPNYYPSKNL 420
QY 417 APIALNAQKFFSTPTIMYNOFLEKTKQLRLDTDOVYGNIAIYFENGRRVRYDTGSN 476
DB 421 APIALNAQDFFSTPTIMN-----YGNIAIYFENGRRVRYDTGSN 460
QY 477 WSEVLPOIETTARIIFNGKDLNVERRIAANVPDPLETTKPDMTLKEALKIARGFNEP 536
DB 461 WSEVLPOIETTARIIFNGKDLNVERRIAANVPDPLETTKPDMTLKEALKIARGFNEP 520
QY 537 NGNLOQOGKDITEFDNFQOOTSQNIKNQLAELNATNIYVLDKIKLNKXNILLIRDKRF 596
DB 521 NGNLOQOGKDITEFDNFQOOTSQNIKNQLAELNATNIYVLDKIKLNKXNILLIRDKRF 580
QY 597 HYDRNNIAGBESVVKAEHREVINSTEGLLLINDKDIRKILSGYIVIEIDTEGLKEVI 656
DB 581 HYDRNNIAGBESVVKAEHREVINSTEGLLLINDKDIRKILSGYIVIEIDTEGLKEVI 640
QY 657 NDRYDMNLSSLRQDGKTFIDFKYNDKPLVYISPNYKVVYAVTKENTIINPSENGDT 716
DB 641 NDRYDMNLSSLRQDGKTFIDFKYNDKPLVYISPNYKVVYAVTKENTIINPSENGDT 700
QY 717 STNGIKKILIFSKKGYEIG 735
DB 701 STNGIKKILIFSKKGYEIG 719

RESULT 9

US-09-273-839A-8
; Sequence 8, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-273-839A-8

Query Match 33.8%; Score 1275; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 9.5e-84;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 486 ETTARIIFNGKDLNVERRIAANVPDPLETTKPDMTLKEALKIARGFNEPNGNLYQOGK 545
DB 29 ETTARIIFNGKDLNVERRIAANVPDPLETTKPDMTLKEALKIARGFNEPNGNLYQOGK 88
QY 546 DITEFDNFQOOTSQNIKNQLAELNATNIYVLDKIKLNKXNILLIRDKRFHYDRNNIAG 605
DB 89 DITEFDNFQOOTSQNIKNQLAELNATNIYVLDKIKLNKXNILLIRDKRFHYDRNNIAG 148
QY 606 GABESVVKAEHREVINSTEGLLLINDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNI 665
DB 149 GABESVVKAEHREVINSTEGLLLINDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNI 208

QY 666 SSIRQDGKTFIDPKYNDKPLVYISPNYKVVYAVTKENTIINPSENGDTSTNGIKKIL 725
DB 209 SSIRQDGKTFIDPKYNDKPLVYISPNYKVVYAVTKENTIINPSENGDTSTNGIKKIL 268
QY 726 IFSKKGYEIG 735
DB 269 IFSKKGYEIG 278

RESULT 10

US-08-960-780-32
; Sequence 32, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Gerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA-708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 177C8
US-08-960-780-32

Query Match 20.3%; Score 768; DB 3; Length 881;
Best Local Similarity 30.1%; Pred. No. 1.1e-46;
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;

QY 4 QENLLNESSSQGLGYFYFSDLNFOAPMVVTSSTTGDLSIPSELEN--IPSENQYFQ 61
DB 39 QKNQ---QKMDRKGLLGYYFKGKDF-SNLTMPAPTRDSTLIYDQQTANKLLDKKQBEYQ 94
QY 62 SAISWGFIVKYSDEYFATSNADHNHTWVDDQEVINKASNSKIRLEKGLYQIKYQ 121

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-31

Query Match 94.6%; Score 3571; DB 1; Length 719;
Best Local Similarity 95.0%; Pred. No. 8.1e-248;
Matches 702; Conservative 4; Mismatches 9; Indels 24; Gaps 2;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPVMTSSITGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSESSQGLGYFSDLNFOAPVMTSSITGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVTSNDNLQPLKQSSNTATIMQGRNFLOQPTVP 176
DB 121 QRENTEKGLDFKLYWTDSONKKEVTSNDNLQPLKQSSNTATIMQGRNFLOQPTVP 176
QY 177 DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPSDF 236
DB 177 DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPSDF 240
QY 237 EKVTRIDKNVSPARPLVAAPFIVHVDMENILSKNEDQSTQNTDSETRISKNTS 296
DB 241 EKVTRIDKNVSPARPLVAAPFIVHVDMENILSKNEDQSTQNTDSETRISKNTS 300
QY 297 RTHTEVHGAENVAHSPFDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGNTAD 356
DB 301 RTHTEVHGAENVAHSPFDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGNTAD 360
QY 357 TARLNARIYNTCTAPINVLPTSLVGNQTLATIDADENQLSOILAPNNYPSKNL 416
DB 361 TARLNARIYNTCTAPINVLPTSLVGNQTLATIDADENQLSOILAPNNYPSKNL 420
QY 417 APIALNAQKFSSTPIITMYNQFLELEKTKQLRLDTPQVYGNATYNFENGRVRVDTGSN 476
DB 421 APIALNAQKFSSTPIITMYNQFLELEKTKQLRLDTPQVYGNATYNFENGRVRVDTGSN 480
QY 477 WSEVLPQIOETARIIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKEALKIATGFNEP 536
DB 461 WSEVLPQIOETARIIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKEALKIATGFNEP 520
QY 537 NGNLQYQKDIETEDFNFDOQTSQNIKNQLAELNATNIYVLDKIKLNACNLIIRDKRF 596
DB 521 NGNLQYQKDIETEDFNFDOQTSQNIKNQLAELNATNIYVLDKIKLNACNLIIRDKRF 580
QY 597 HYDRNNIAGDAESVYKAEHREVNINSSTEGLLNIDKIDKILSGYIIEDETEGLKEVI 656
DB 581 HYDRNNIAGDAESVYKAEHREVNINSSTEGLLNIDKIDKILSGYIIEDETEGLKEVI 640

657 NDRYDMLNISSLRQDGKTFIDFKKNDKULPLVISHPNYKVNKYAVTKENTINPSENGT 716
DB 641 NDRYDMLNISSLRQDGKTFIDFKKNDKULPLVISHPNYKVNKYAVTKENTINPSENGT 700
QY 717 STNGIKKILIFSKGYEIG 735
DB 701 STNGIKKILIFSKGYEIG 719

RESULT 8
PCT-US94-01624-31
Sequence 31, Application PC/TUS9401624
GENERAL INFORMATION:
APPLICANT: Leppla, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
TITLE OF INVENTION: RELATED METHODS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
STREET: Steuart Street Tower, 20th Floor, One Market
STREET: Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01624
FILING DATE: June 25, 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-31

Query Match 94.6%; Score 3571; DB 5; Length 719;
Best Local Similarity 95.0%; Pred. No. 8.1e-248;
Matches 702; Conservative 4; Mismatches 9; Indels 24; Gaps 2;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPVMTSSITGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSESSQGLGYFSDLNFOAPVMTSSITGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVTSNDNLQPLKQSSNTATIMQGRNFLOQPTVP 176
DB 121 QRENTEKGLDFKLYWTDSONKKEVTSNDNLQPLKQSSNTATIMQGRNFLOQPTVP 180
QY 177 DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPSDF 236
DB 181 DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPSDF 240

Db 421 LNAQDDFSSTPTMNVNQFLEKTKQLRLDQVYGNIAVYVNGRVRVDTGSMSEV 480
 QY 481 LPQIQTETARIIFNGKDLNVLVRRIRAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
 Db 481 LPQIQTETARIIFNGKDLNVLVRRIRAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
 QY 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNITVLDKIKLNKXNMLIRDKRFHYDR 600
 Db 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNITVLDKIKLNKXNMLIRDKRFHYDR 600
 QY 601 NNIAGADESVVKEAHREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
 Db 601 NNIAGADESVVKEAHREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
 QY 661 DMLNSSLRQDQKTFIDFKKYNDKPLXIYISNPNKVNVAVTYKNTIINPSENGDTSTNG 720
 Db 661 DMLNSSLRQDQKTFIDFKKYNDKPLXIYISNPNKVNVAVTYKNTIINPSENGDTSTNG 720
 QY 721 IKKIL---IFSCKG 731
 Db 721 IKKILKKVILGKG 734

RESULT 6
 PCT-US94-01624-12
 ; Sequence 12, Application PC/TUS9401624
 ; GENERAL INFORMATION:
 ; APPLICANT: Leppla, Stephen H.
 ; APPLICANT: Klimpel, Kurt R.
 ; APPLICANT: Arora, Naveen
 ; APPLICANT: Singh, Yogendra
 ; APPLICANT: Nichols, Peter J.
 ; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
 ; TITLE OF INVENTION: RELATED METHODS
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW
 ; STREET: Steuart Street Tower, 20th Floor, One Market
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/01624
 ; FILING DATE: June 25, 1993
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Kenneth A.
 ; REGISTRATION NUMBER: 31,677
 ; REFERENCE/DOCKET NUMBER: 15280-115
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 543-9600
 ; TELEFAX: (415) 543-5043
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 903 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US94-01624-12

Query Match 98.1%; Score 3703.5; DB 5; Length 903;
 Best Local Similarity 98.6%; Pred. No. 3.5e-257;
 Matches 724; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 EVKQENLLNESSSQGLLYVFDLNFQAPMVVTSSTTGDLSIPSSLENIPSENQYF 60

Db 1 EVKQENLLNESSSQGLLYVFDLNFQAPMVVTSSTTGDLSIPSSLENIPSENQYF 60
 QY 61 QSAIWGSGFIKKVKSDEYTFATSDADNHVMTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
 Db 61 QSAIWGSGFIKKVKSDEYTFATSDADNHVMTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
 QY 121 QRENPTKEGLDFKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKKRSTAGPTVPDRDN 180
 Db 121 QRENPTKEGLDFKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKKRSTAGPTVPDRDN 180
 QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDPEKVT 240
 Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDPEKVT 240
 QY 241 GRIDKNVSPARPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTTISKNTSTSRHT 300
 Db 241 GRIDKNVSPARPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTTISKNTSTSRHT 300
 QY 301 SEVHNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
 Db 301 SEVHNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
 QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPPSKNLAPIA 420
 Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPPSKNLAPIA 420
 QY 421 LNAQKFSSTPTMNVNQFLEKTKQLRLDQVYGNIAVYVNGRVRVDTGSMSEV 480
 Db 421 LNAQKFSSTPTMNVNQFLEKTKQLRLDQVYGNIAVYVNGRVRVDTGSMSEV 480
 QY 481 LPQIQTETARIIFNGKDLNVLVRRIRAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
 Db 481 LPQIQTETARIIFNGKDLNVLVRRIRAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
 QY 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNITVLDKIKLNKXNMLIRDKRFHYDR 600
 Db 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNITVLDKIKLNKXNMLIRDKRFHYDR 600
 QY 601 NNIAGADESVVKEAHREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
 Db 601 NNIAGADESVVKEAHREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
 QY 661 DMLNSSLRQDQKTFIDFKKYNDKPLXIYISNPNKVNVAVTYKNTIINPSENGDTSTNG 720
 Db 661 DMLNSSLRQDQKTFIDFKKYNDKPLXIYISNPNKVNVAVTYKNTIINPSENGDTSTNG 720
 QY 721 IKKIL---IFSCKG 731
 Db 721 IKKILKKVILGKG 734

RESULT 7
 US-08-082-849B-31
 ; Sequence 31, Application US/08082849B
 ; Patent No. 5677274
 ; GENERAL INFORMATION:
 ; APPLICANT: Leppla, Stephen H.
 ; APPLICANT: Klimpel, Kurt R.
 ; APPLICANT: Arora, Naveen
 ; APPLICANT: Singh, Yogendra
 ; APPLICANT: Nichols, Peter J.
 ; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
 ; TITLE OF INVENTION: Related Methods
 ; NUMBER OF SEQUENCES: 35
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

```
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 903 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-021-601-12

Query Match      98.1%; Score 3703.5; DB 1; Length 903;
Best Local Similarity 98.6%; Pred. No. 3.5e-257;
Matches 724; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
Db 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
QY 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVDRDN 180
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKWKSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFFSPTITWYVNTQFLEKTKQLRLDQVYGNATYNPENGRVVDTSNNSSEV 480
Db 421 LNAQKFFSPTITWYVNTQFLEKTKQLRLDQVYGNATYNPENGRVVDTSNNSSEV 480

Query Match      98.1%; Score 3703.5; DB 1; Length 903;
Best Local Similarity 98.6%; Pred. No. 3.5e-257;
Matches 724; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
Db 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
QY 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVDRDN 180
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKWKSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFFSPTITWYVNTQFLEKTKQLRLDQVYGNATYNPENGRVVDTSNNSSEV 480
Db 421 LNAQKFFSPTITWYVNTQFLEKTKQLRLDQVYGNATYNPENGRVVDTSNNSSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540
QY 541 QYQKGDITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKQNILIRDKRPHYDR 600
Db 541 QYQKGDITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKQNILIRDKRPHYDR 600
QY 601 NNIAVGADESVEKAEHREVINSSTEGILLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVEKAEHREVINSSTEGILLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKYVNVYATKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNKYVNVYATKENTIIINPSENGDTSTNG 720
QY 721 IKKIL---IFSKG 731
Db 721 IKKILKKVWLGKG 734

RESULT 5
US-08-082-849B-12
; Sequence 12, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Aroa, Naveen
```

QY 661 DMLNSSLRQDQKTFIDFKKNDKLPXYISNPYKVNVAVTKENTINPSENGDTSTNG 720
DB 661 DMLNSSLRQDQKTFIDFKKNDKLPXYISNPYKVNVAVTKENTINPSENGDTSTNG 720
QY 721 IKKILFSSKGYEIG 735
DB 721 IKKILFSSKGYEIG 735
RESULT 3
PCT-US94-01624-4
; Sequence 4, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE and CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01624-4
Query Match 99.3%; Score 3748; DB 5; Length 735;
Best Local Similarity 99.5%; Pred. No. 1.7e-260; Mismatches 4; Indels 0; Gaps 0;
Matches 731; Conservative 0;
QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMWTSTTGDLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNSESQGLGYFSDLNFOAPMWTSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWGFFIKVKSDEYVTSATSDNHTVMWDDQEVINKASNKIRLEKGLYQIKQY 120
DB 61 QSAIWGFFIKVKSDEYVTSATSDNHTVMWDDQEVINKASNKIRLEKGLYQIKQY 120
QY 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNQLQPELKQSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNQLQPELKQSSNRKRSSTAGTVPDRDN 180
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSPKRWSTASDPSDFEYV 240
DB 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSPKRWSTASDPSDFEYV 240

QY 241 GRIDKNVSPARHPLVAAYPVHVDMENIILSKNEQSQNTDSETRTISKNVTSRTH 300
DB 241 GRIDKNVSPARHPLVAAYPVHVDMENIILSKNEQSQNTDSETRTISKNVTSRTH 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNNVLTPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYNNVLTPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTMTYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSKNWEV 480
DB 421 LNAQKFSSTPTMTYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSKNWEV 480
QY 481 LPOIQETTARIIFNGKDLNLVRRRIAANVPSDPLETTKPDWTLKEALKIAGFNEPENG 540
DB 481 LPOIQETTARIIFNGKDLNLVRRRIAANVPSDPLETTKPDWTLKEALKIAGFNEPENG 540
QY 541 QYQKDIETPFDNFDOQTSONIKVQLAELNATNIYTVLDKIKLNAKONILIRDKRPHYDR 600
DB 541 QYQKDIETPFDNFDOQTSONIKVQLAELNATNIYTVLDKIKLNAKONILIRDKRPHYDR 600
QY 601 NNTAVGADESVMKEAHEVINSSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660
DB 601 NNTAVGADESVMKEAHEVINSSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDQKTFIDFKKNDKLPXYISNPYKVNVAVTKENTINPSENGDTSTNG 720
DB 661 DMLNSSLRQDQKTFIDFKKNDKLPXYISNPYKVNVAVTKENTINPSENGDTSTNG 720
QY 721 IKKILFSSKGYEIG 735
DB 721 IKKILFSSKGYEIG 735
RESULT 4
US-08-021-601-12
; Sequence 12, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 133 Carnegie Way, Suite 400
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/021,601
; FILING DATE: 19930212
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880

1 EVKQENRLNSESSESSQGLLYFFSLNFPQAPMVVTSSTTGDLSIPSELENIPSENOYF 60
61 QSAIWSGFTKVKKSDEYTFATSDADNHVTWVDDQVINKANSNKIRLEKGRLYQIKIY 120
61 QSAIWSGFTKVKKSDEYTFATSDADNHVTWVDDQVINKANSNKIRLEKGRLYQIKIY 120
121 QRENTEKGLDFKLYWTDSONKKEVSSNLOLPELKOKSNSRKRSTSGPTVPDRDN 180
121 QRENTEKGLDFKLYWTDSONKKEVSSNLOLPELKOKSNSRKRSTSGPTVPDRDN 180
181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKT 240
181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKT 240
241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
301 SEVHGNAEVHASFDDIGGSVSGFSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360
301 SEVHGNAEVHASFDDIGGSVSGFSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360
361 NANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
361 NANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
421 LNAQKPSSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
421 LNAQKPSSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540
481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540
541 QYQKDIETEFDFPDQTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600
541 QYQKDIETEFDFPDQTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600
601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

RESULT 2

US-08-082-849B-4
Sequence 4, Application US/08082849B
Patent No. 567274

GENERAL INFORMATION:

APPLICANT: Leppla, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
TITLE OF INVENTION: Related Methods

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-4

Query Match 99.3%; Score 3748; DB 1; Length 735;

Best Local Similarity 99.5%; Pred. No. 1.7e-260;

Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFFSLNFPQAPMVVTSSTTGDLSIPSELENIPSENOYF 60
DB 1 EVKQENRLNSESSESSQGLLYFFSLNFPQAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFTKVKKSDEYTFATSDADNHVTWVDDQVINKANSNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFTKVKKSDEYTFATSDADNHVTWVDDQVINKANSNKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSNLOLPELKOKSNSRKRSTSGPTVPDRDN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVSSNLOLPELKOKSNSRKRSTSGPTVPDRDN 180
QY 181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKT 240
DB 181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSGFSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSGFSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKPSSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
DB 421 LNAQKPSSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
QY 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540
DB 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540
QY 541 QYQKDIETEFDFPDQTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600
DB 541 QYQKDIETEFDFPDQTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:31:43 ; Search time 15.0986 Seconds
(without alignments)
2513.152 Million cell updates/sec

Title: US-09-848-909A-11
Perfect score: 3774
Sequence: 1 EVKQENLLNSESQGLL.....TSTNGIKLIFSKGVEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 5125971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PTCUS-COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result: being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3748	99.3	735	1 US-08-021-601-4	Sequence 4, Appl
2	3748	99.3	735	1 US-08-082-849B-4	Sequence 4, Appl
3	3748	99.3	735	5 PCT-US94-01624-4	Sequence 4, Appl
4	3703.5	98.1	903	1 US-08-021-601-12	Sequence 12, Appl
5	3703.5	98.1	903	1 US-08-082-849B-12	Sequence 12, Appl
6	3703.5	98.1	903	5 PCT-US94-01624-12	Sequence 12, Appl
7	3571	94.6	719	1 US-08-082-849B-31	Sequence 31, Appl
8	3571	94.6	719	5 PCT-US94-01624-31	Sequence 31, Appl
9	1275	33.8	288	4 US-09-273-839A-8	Sequence 8, Appl
10	768	20.3	881	3 US-08-960-780-32	Sequence 32, Appl
11	768	20.3	881	3 US-09-073-898-32	Sequence 32, Appl
12	768	20.3	881	4 US-09-307-106-8	Sequence 8, Appl
13	768	20.3	881	4 US-09-850-351A-32	Sequence 32, Appl
14	766	20.3	884	1 US-08-471-033-5	Sequence 5, Appl
15	766	20.3	884	2 US-08-471-044-5	Sequence 5, Appl
16	766	20.3	884	2 US-08-463-483A-5	Sequence 5, Appl
17	766	20.3	884	2 US-08-471-046A-5	Sequence 5, Appl
18	766	20.3	884	2 US-08-470-566B-5	Sequence 5, Appl
19	766	20.3	884	2 US-08-469-334-5	Sequence 5, Appl
20	766	20.3	884	3 US-09-300-529-5	Sequence 5, Appl
21	766	20.3	1346	1 US-08-471-033-23	Sequence 23, Appl
22	766	20.3	1346	2 US-08-471-044-23	Sequence 23, Appl
23	766	20.3	1346	2 US-08-463-483A-23	Sequence 23, Appl
24	766	20.3	1346	2 US-08-471-046A-23	Sequence 23, Appl
25	766	20.3	1346	2 US-08-470-566B-23	Sequence 23, Appl
26	766	20.3	1346	2 US-08-469-334-23	Sequence 23, Appl
27	766	20.3	1346	3 US-09-300-529-23	Sequence 23, Appl

28	765	20.3	852	1 US-08-471-033-36	Sequence 36, Appl
29	765	20.3	852	2 US-08-471-044-36	Sequence 36, Appl
30	765	20.3	852	2 US-08-463-483A-36	Sequence 36, Appl
31	765	20.3	852	2 US-08-471-046A-36	Sequence 36, Appl
32	765	20.3	852	2 US-08-470-566B-36	Sequence 36, Appl
33	765	20.3	852	2 US-08-469-334-36	Sequence 36, Appl
34	765	20.3	852	3 US-09-300-529-36	Sequence 36, Appl
35	765	20.3	1338	1 US-08-471-033-50	Sequence 50, Appl
36	765	20.3	1338	2 US-08-471-044-50	Sequence 50, Appl
37	765	20.3	1338	2 US-08-463-483A-50	Sequence 50, Appl
38	765	20.3	1338	2 US-08-471-046A-50	Sequence 50, Appl
39	765	20.3	1338	2 US-08-470-566B-50	Sequence 50, Appl
40	765	20.3	1338	2 US-08-469-334-50	Sequence 50, Appl
41	765	20.3	1338	3 US-09-300-529-50	Sequence 50, Appl
42	755.5	20.0	784	3 US-09-371-913A-7	Sequence 7, Appl
43	755.5	20.0	784	4 US-09-967-803-7	Sequence 7, Appl
44	746.5	19.8	860	4 US-09-307-106-48	Sequence 48, Appl
45	736.5	19.5	834	1 US-08-471-033-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-021-601-4
; Sequence 4, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Arcora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 133 Carnegie Way, Suite 400
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/021,601
; FILING DATE: 19930212
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/POCKET NUMBER: 1414.057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-9770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-021-601-4

Query Match 99.3%; Score 3748; DB 1; Length 735;
Best Local Similarity 99.5%; Pred. No. 1.7e-260;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENLLNSESQGLLGYFSDNLFQAPMVVTSSTTCDDLSIPSELENISENQYF 60
|||||

DB 690 DMLNSSLRQDGKTFIDFKYNDKLPFLYISNPYKVVVAVTKENTIIINPSENGDTSTNG 749

QY 721 IKKILIFSKKGYEIG 735

DB 750 IKKILIFSKKGYEIG 764

RESULT 15

AAB47306

ID AAB47306 standard; protein; 764 AA.

AC AAB47306;

DT 29-AUG-2001 (first entry)

DE Wild type B. anthracis protective antigen.

XX Lethal factor; LF; immunogen; LF4; protective antigen; PA; DNA vaccine;

KW humoral; cell-mediated; immune memory response.

XX Bacillus anthracis.

XX Key Location/Qualifiers

FT Peptide 1..29

FT /label= signal peptide

FT /note= "Not given in the specification"

FT Protein 30..764

FT /label= PA

FT Peptide 204..764

FT /label= pCPA

XX WO2000145639-A2.

PN 28-JUN-2001.

XX 21-DEC-2000; 2000WO-US034912.

PF 22-DEC-1999; 99US-0171459P.

PR (OHIS) UNIV OHIO STATE RES FOUND.

XX (GALL/) GALLOWAY D R.

PA (MATE/) MATECZUN A J.

XX Galloway DR, Mateczun AJ;

PI WPI; 2001-408540/43.

XX N-PSDB; AAC86016.

DR Protecting animal against lethal infection with Bacillus anthracis, by

XX administering wildtype or mutated form of Bacillus anthracis lethal

PT factor protein or its fragment or a nucleic acid encoding the mutated

PT protein.

XX Claim 5; Fig 2; 33pp; English.

PS This sequence shows the B. anthracis protective antigen (PA). An

CC immunogenic fragment of PA, pCPA, can be used to produce an immune

CC response which protects an animal against lethal infection with Bacillus

CC anthracis. DNA encoding the B. anthracis PA can be used in conjunction

CC with DNA encoding the lethal factor (LF) in a DNA vaccine. Using a DNA

CC vaccine which encodes the mutated LF protein or fragment alone or in

CC combination with a DNA encoding the PA protein or its fragment, both

CC components (humoral and cell-mediated) of the immune system are

CC stimulated, which results in longer term immune memory response. The

CC combined use of a mutated LF and PA gene or their fragments results in a

CC higher level of immune response as judged by overall serum antibody in

CC titers for LF and PA antigens, than the use of either LF or PA genes in

CC separate immunizations

XX Sequence 764 AA;

Query Match 99.3%; Score 3748; DB 4; Length 764;

Best Local Similarity 99.5%; Pred. No. 1.3e-239;

Matches	731;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
QY	1	EYKQENRLNSESSESSQGLGYPFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENQYF	60						
DB	30	EYKQENRLNSESSESSQGLGYPFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENQYF	89						
QY	61	QSAIWSGFIKVKKSDVENTFATSADNHVTMMVDDQEVINKASNNKIRLEKGRLYQIKIY	120						
DB	90	QSAIWSGFIKVKKSDVENTFATSADNHVTMMVDDQEVINKASNNKIRLEKGRLYQIKIY	149						
QY	121	QRENPTKGLDFKLYWTDSONKKEVISDNILQLPKQKSNRKKKSTSGAGTVPDRDN	180						
DB	150	QRENPTKGLDFKLYWTDSONKKEVISDNILQLPKQKSNRKKKSTSGAGTVPDRDN	209						
QY	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKMTAGDPVSDPEKVT	240						
DB	210	DGIPDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKMTAGDPVSDPEKVT	269						
QY	241	GRIDKNVSPPEARHPLVAAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKTSTSTHT	300						
DB	270	GRIDKNVSPPEARHPLVAAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKTSTSTHT	329						
QY	301	SEVHGNAEVHASFDPDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITWGLNTADTARL	360						
DB	330	SEVHGNAEVHASFDPDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITWGLNTADTARL	389						
QY	361	NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPPSKNLAPIA	420						
DB	390	NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPPSKNLAPIA	449						
QY	421	LNAOKKFSSTPITMNYNQLFLEKTKQLRLDTDOVYGNATYFENGRVAVDTGSKNWEV	480						
DB	450	LNAOKKFSSTPITMNYNQLFLEKTKQLRLDTDOVYGNATYFENGRVAVDTGSKNWEV	509						
QY	481	LPQIETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAPGNEPNGL	540						
DB	510	LPQIETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAPGNEPNGL	569						
QY	541	OYCGKDIETFDNFDOQTSONIKQLAELNATNYTVLDKIKNAXKNILLIRDKRFHYDR	600						
DB	570	OYCGKDIETFDNFDOQTSONIKQLAELNATNYTVLDKIKNAXKNILLIRDKRFHYDR	629						
QY	601	NNIAYGADESIVKEAHREVINSSTEGLLNLDKIRKILSGYIIVEIEDTEGLKEVINDRY	660						
DB	630	NNIAYGADESIVKEAHREVINSSTEGLLNLDKIRKILSGYIIVEIEDTEGLKEVINDRY	689						
QY	661	DMLNSSLRQDGKTFIDFKYNDKLPFLYISNPYKVVVAVTKENTIIINPSENGDTSTNG	720						
DB	690	DMLNSSLRQDGKTFIDFKYNDKLPFLYISNPYKVVVAVTKENTIIINPSENGDTSTNG	749						
QY	721	IKKILIFSKKGYEIG 735							
DB	750	IKKILIFSKKGYEIG 764							

Search completed: May 3, 2004, 19:36:00

Job time : 51.1208 secs

```
QY 121 QRENPTKGLDFKLYWTDSONKKEVISNDNLQLPDLKOKSSNRKKESTAGTVPDRDN 180
Db 149 QRENPTKGLDFKLYWTDSONKKEVISNDNLQLPDLKOKSSNRKKESTAGTVPDRDN 208
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKNSASTADPYSDPEKVT 240
Db 209 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKNSASTADPYSDPEKVT 268
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 300
Db 269 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 328
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 329 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 388
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQSLQILAPNNYPPSKNLAPIA 420
Db 389 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQSLQILAPNNYPPSKNLAPIA 448
QY 421 LNAOKKRSSTPTIMNYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 480
Db 449 LNAQDDFSTPTIMNYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 508
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL 540
Db 509 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL 568
QY 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNNILIRDKRFHYDR 600
Db 569 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNNILIRDKRFHYDR 628
QY 601 NNIAVGADES VVKEAAREVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
Db 629 NNIAVGADES VVKEAAREVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 688
QY 661 DMLNITSSLRQDGTFFIDFKKYNKPLIYISNPYKVNVAVTKENTIIINPSENGDTSNG 720
Db 689 DMLNITSSLRQDGTFFIDFKKYNKPLIYISNPYKVNVAVTKENTIIINPSENGDTSNG 748
QY 721 IKKILIFSKKGYEIG 735
Db 749 IKKILIFSKKGYEIG 763

RESULT 14
AA56958
ID AA56958 standard; protein; 764 AA.
XX
AC AA56958;
XX
DT 25-APR-2000 (first entry)
XX
DE B. anthracis protective antigen (PA) protein.
XX
KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;
KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.
XX
OS Bacillus anthracis.
XX
PN WO200002522-A2.
XX
PD 20-JAN-2000.
XX
PF 09-JUL-1999; 99WO-US015568.
XX
PR 10-JUL-1998; 98US-0092416P.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;
XX
```

```
DR WPI; 2000-182165/16.
DR N-PSDB; AA256874.
XX
PT Recombinant DNA construct useful as vaccines for anthrax, in producing
PT host cells for analyzing the drugs and agents inhibiting anthrax.
XX
PS Disclosure; Page 33; 35pp; English.
XX
CC The invention provides a recombinant DNA construct that comprises a
CC vector and at least one nucleic acid (or its fragment) encoding a
CC combination of Bacillus anthracis proteins, selected from protective
CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA
CC with its secretory signals replaced with those of tissue plasminogen
CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine
CC for anthrax and in producing infectious alpha virus particles. These
CC particles, expressing the B. anthracis proteins are useful also as
CC vaccines for anthrax. Host cells transformed with the construct are
CC useful for analyzing the effectiveness of drugs and agents that inhibit
CC anthrax or B. anthracis proteins. The present sequence represents a B.
CC anthracis PA protein
XX
SQ Sequence 764 AA;
Query Match 99.3%; Score 3748; DB 3; Length 764;
Best Local Similarity 99.5%; Pred. No. 1.3e-239;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 EVKQENRLNLESSESSQGLLYFFDLNFAQPMVVTSSTTGDLSPSSSELENIPISENQVF 60
Db 30 EVKQENRLNLESSESSQGLLYFFDLNFAQPMVVTSSTTGDLSPSSSELENIPISENQVF 89
QY 61 QSAIWSGFTKVKKSDSYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 90 QSAIWSGFTKVKKSDSYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 149
QY 121 QRENPTKGLDFKLYWTDSONKKEVISNDNLQLPDLKOKSSNRKKESTAGTVPDRDN 180
Db 150 QRENPTKGLDFKLYWTDSONKKEVISNDNLQLPDLKOKSSNRKKESTAGTVPDRDN 209
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKNSASTADPYSDPEKVT 240
Db 210 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKNSASTADPYSDPEKVT 269
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 300
Db 270 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 329
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 330 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 389
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQSLQILAPNNYPPSKNLAPIA 420
Db 390 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQSLQILAPNNYPPSKNLAPIA 449
QY 421 LNAOKKRSSTPTIMNYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 480
Db 450 LNAQDDFSTPTIMNYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 509
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL 540
Db 510 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL 569
QY 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNNILIRDKRFHYDR 600
Db 570 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNNILIRDKRFHYDR 629
QY 601 NNIAVGADES VVKEAAREVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
Db 630 NNIAVGADES VVKEAAREVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 689
QY 661 DMLNITSSLRQDGTFFIDFKKYNKPLIYISNPYKVNVAVTKENTIIINPSENGDTSNG 720
```


XX	(USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX	Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;
XX	WPI; 2000-182165/16.
XX	N-PsDB; NAAZ56875.
DR	Recombinant DNA construct useful as vaccines for anthrax, in producing
XX	host cells for analyzing the drugs and agents inhibiting anthrax.
XX	Disclosure; Page 34; 35pp; English.
XX	The invention provides a recombinant DNA construct that comprises a
XX	vector and at least one nucleic acid (or its fragment) encoding a
CC	combination of Bacillus anthracis proteins, selected from protective
CC	antigen (PA) MAT-PA (PA with its secretory signals removed), tPA-PA (PA
CC	with its secretory signals replaced with those of tissue plasminogen
CC	activator) and PA63 (63 KDa PA). The construct is useful as a DNA vaccine
CC	for anthrax and in producing infectious alpha virus particles. These
CC	particles, expressing the B. anthracis proteins are useful also as
CC	vaccines for anthrax. Host cells transformed with the construct are
CC	useful for analyzing the effectiveness of drugs and agents that inhibit
CC	anthrax or B. anthracis proteins. The present sequence represents a B.
CC	anthracis MAT-PA protein
XX	Sequence 736 AA;
XX	Query Match 99.3%; Score 3748; DB 3; Length 736;
CC	Best Local Similarity 99.5%; Pred. No. 1.2e-239;
XX	Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	1 EVKQENRLNSESSSQGLGYFSDLNFOAPWVTSITTTGDLSPSSELENIPISENQYF 60
DB	2 EVKQENRLNSESSSQGLGYFSDLNFOAPWVTSITTTGDLSPSSELENIPISENQYF 61
QY	61 QSAIWSGFIKYKSDSEYTFATSAADNHVTMWDDQEVINKASNSNKIRLEKGRLYQIKIY 120
DB	62 QSAIWSGFIKYKSDSEYTFATSAADNHVTMWDDQEVINKASNSNKIRLEKGRLYQIKIY 121
QY	121 QRENTEKGLDFKLYWTDSONKQDEVISSDNLQLPELKQKSSNRKRSSTAGTVPDRDN 180
DB	122 QRENTEKGLDFKLYWTDSONKQKEVISSDNLQLPELKQKSSNRKRSSTAGTVPDRDN 181
QY	181 DGIPOSLEVGTVVQKRRITFLSPWISNHEKGLTKYKSSPEKWSSTADPYSDPEKVT 240
DB	182 DGIPOSLEVGTVVQKRRITFLSPWISNHEKGLTKYKSSPEKWSSTADPYSDPEKVT 241
QY	241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB	242 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 301
QY	301 SEVHGNAEVHSAFPDGGSVGAGSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
DB	302 SEVHGNAEVHSAFPDGGSVGAGSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 361
QY	361 NANIRVYNTGTAPIYVLPPTTSLVLGKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB	362 NANIRVYNTGTAPIYVLPPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 421
QY	421 LNAQKFSSTPTIMYNYQFLEKTKQIRLDTQVYGNITATYAFENGSRVVDGTSNWSEV 480
DB	422 LNAQDDFSSTPTIMYNYQFLEKTKQIRLDTQVYGNITATYAFENGSRVVDGTSNWSEV 481
QY	481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 540
DB	482 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 541
QY	541 QYQKGDITFDNFDOOTSQNLKQALAEINATNYITVLDDKIKLNKSNKILIRDKRPHYDR 600
DB	542 QYQKGDITFDNFDOOTSQNLKQALAEINATNYITVLDDKIKLNKSNKILIRDKRPHYDR 601
QY	601 NNTAVGADESVVKEAHREVINSSSTGLLNLDKQIRKILSGYIVIEIDTTLKEVINDRY 660

```

Db 541 QYQKDIETDFDFNQTSQNIKNQLAELNATNYITVLDKIKNAKNMILIRDRFHYDR 600
QY 601 NNIAVGADESVMKEAAREVINSSTGLINIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAAREVINSSTGLINIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRODQKTFIDFKYNDKULPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODQKTFIDFKYNDKULPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 11
AAM51488
ID AAM51488 standard; protein; 735 AA.
AC AAM51488;
XX
DT 01-FEB-2002 (first entry)
DE Anthrax PA mutant D425A.
XX
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX
OS Bacillus anthracis.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 425 /note= "Wild-type Asp substituted by Ala"
FT
XX
FN WC200182788-A2.
XX
XX
XX 08-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014372.
XX
XX 04-MAY-2000; 2000US-0201800P.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX
XX Claim 4; Page; 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
XX Sequence 735 AA;
XX
XX Query Match 99.3%; Score 3748; DB 5; Length 735;
XX Best Local Similarity 99.5%; Pred. No. 1.2e-239;
XX Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 1 EVKQENRLNSESQGLLYFFDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60

```

```

Db 1 EVKQENRLNSESQGLLYFFDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFLYKLTDSQNKKEVSSDNQLPELKQKSSNSRKKRSTAGTVPDRDN 180
Db 121 QRENPTKGLDFLYKLTDSQNKKEVSSDNQLPELKQKSSNSRKKRSTAGTVPDRDN 180
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPPYDSEKVT 240
Db 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPPYDSEKVT 240
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKVEDQSTQNTDSETRTISKNTSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKVEDQSTQNTDSETRTISKNTSTRTHT 300
QY 301 SEVHGNAEVSHPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITMGLTADTARL 360
Db 301 SEVHGNAEVSHPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITMGLTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSOLAPNNYPPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSOLAPNNYPPSKNLAPIA 420
QY 421 LNAQKFSSTPIITMNYNOFLEKTKQLRLDQVYGNATYNFENGVRVVDVTGSNWSEV 480
Db 421 LNAQKFSSTPIITMNYNOFLEKTKQLRLDQVYGNATYNFENGVRVVDVTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNIVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENGNL 540
Db 481 LPOIQTETARIIFNGKDLNIVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENGNL 540
QY 541 QYQKDIETDFDFNQTSQNIKNQLAELNATNYITVLDKIKNAKNMILIRDRFHYDR 600
Db 541 QYQKDIETDFDFNQTSQNIKNQLAELNATNYITVLDKIKNAKNMILIRDRFHYDR 600
QY 601 NNIAVGADESVMKEAAREVINSSTGLINIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAAREVINSSTGLINIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRODQKTFIDFKYNDKULPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODQKTFIDFKYNDKULPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 12
AAY56959
ID AAY56959 standard; protein; 736 AA.
XX
XX AC AAY56959;
XX
XX DT 25-APR-2000 (first entry)
XX
XX DE B. anthracis MAT-PA protein.
XX
XX KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;
XX tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.
XX
XX OS Bacillus anthracis.
XX
XX FN WO200002522-A2.
XX
XX PD 20-JAN-2000.
XX
XX PF 09-JUL-1999; 99WO-US015568.
XX
XX 10-JUL-1998; 98US-0092416P.
XX

```

```

SQ      Sequence 735 AA;
Query Match      99.3%; Score 3748; DB 2; Length 735;
Best Local Similarity 99.5%; Pred. No. 1.2e-239;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60
DB      1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60

QY      61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
DB      61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120

QY      121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPDLKQKSSNSRKKRSTAGPTVPDRN 180
DB      121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPDLKQKSSNSRKKRSTAGPTVPDRN 180

QY      181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYV 240
DB      181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYV 240

QY      241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB      241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY      301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
DB      301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAGTGLNTADTARL 360

QY      361 NANIRYVNTGTAPIYVNLPTTSLVGNKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB      361 NANIRYVNTGTAPIYVNLPTTSLVGNKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420

QY      421 LNAQKFFSTPTITMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480
DB      421 LNAQKFFSTPTITMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480

QY      481 LPOIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB      481 LPOIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

QY      541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRFHYDR 600
DB      541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRFHYDR 600

QY      601 NNIAVGADESVMKEAHREVINSGTGLNINIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
DB      601 NNIAVGADESVMKEAHREVINSGTGLNINIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660

QY      661 DMLNISSLFQDGKTFIDPKKYNDKPLXIISNPNYKVNYAVTKENTIINPSENGDTSNG 720
DB      661 DMLNISSLFQDGKTFIDPKKYNDKPLXIISNPNYKVNYAVTKENTIINPSENGDTSNG 720

QY      721 IKKILIFSCKGYEIG 735
DB      721 IKKILIFSCKGYEIG 735

RESULT 10
AAM51483
ID      AAM51483 standard; protein; 735 AA.
XX
AC      AAM51483;
XX
DT      01-FEB-2002 (first entry)
XX
DE      Anthrax PA protein.
XX
KW      Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX      B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.
XX
OS      Bacillus anthracis.
```

```

XX      WO200182788-A2.
XX
XX      08-NOV-2001.
XX
XX      04-MAY-2001; 2001WO-US014372.
XX
XX      04-MAY-2000; 2000US-0201800P.
XX
XX      (HARD ) HARVARD COLLEGE.
XX
XX      Collier RJ, Sellman BR;
XX
XX      MPI; 2002-017725/02.
XX
XX      N-PSDB; AA:99904.
XX
XX      Protecting humans against anthrax using mutant B groups (anthrax
XX      protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX      anthracis.
XX
XX      Disclosure; Fig 13; 77pp; English.
XX
XX      The invention relates to antibacterial agents comprising mutant forms of
XX      pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX      in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX      moiety is anthrax protective antigen (PA) and using these mutants or
XX      compositions of them for protecting against Bacillus anthracis infections
XX      in humans, especially as vaccines. The present sequence is that of the
XX      anthrax PA protein
XX
XX      Sequence 735 AA;
Query Match      99.3%; Score 3748; DB 5; Length 735;
Best Local Similarity 99.5%; Pred. No. 1.2e-239;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60
DB      1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60

QY      61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
DB      61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120

QY      121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPDLKQKSSNSRKKRSTAGPTVPDRN 180
DB      121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPDLKQKSSNSRKKRSTAGPTVPDRN 180

QY      181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYV 240
DB      181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYV 240

QY      241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB      241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY      301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
DB      301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAGTGLNTADTARL 360

QY      361 NANIRYVNTGTAPIYVNLPTTSLVGNKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB      361 NANIRYVNTGTAPIYVNLPTTSLVGNKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420

QY      421 LNAQKFFSTPTITMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480
DB      421 LNAQKFFSTPTITMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480

QY      481 LPOIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB      481 LPOIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

QY      541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRFHYDR 600
DB      541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRFHYDR 600
```

KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
 XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
 OS Bacillus anthracis.
 XX Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 425 /note= "Wild-type Asp substituted by Asn"
 XX WO200182788-A2.
 XX PD 08-NOV-2001.
 XX PF 04-MAY-2001; 2001WO-US014372.
 XX PR 04-MAY-2000; 2000US-0201800P.
 XX PA (HARD) HARVARD COLLEGE.
 XX PI Collier RJ, Sellman BR;
 XX DR WPI; 2002-017725/02.
 XX PT Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.
 XX Claim 4; Page; 77pp; English.
 XX The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections
 CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)
 XX SQ Sequence 735 AA;

Query Match 99.3%; Score 3749; DB 5; Length 735;
 Best Local Similarity 99.5%; Pred. No. 1.1e-239;
 Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNQFQPMVVTSTTGDLSIPSELENIPSENYF 60
 DB 1 EVKQENRLNSESSESSQGLGYFSDLNQFQPMVVTSTTGDLSIPSELENIPSENYF 60
 QY 61 QSAIWSGFIKVKSDYTFATSAADNHVTMWYDDQEVINKASNSKIRLEKGRLYQIKIY 120
 DB 61 QSAIWSGFIKVKSDYTFATSAADNHVTMWYDDQEVINKASNSKIRLEKGRLYQIKIY 120
 QY 121 QRENTEKGLDPLKLYWTDSONKKEVISSDNQLQPLKQKSNRSKSTSGAPVPRDN 180
 DB 121 QRENTEKGLDPLKLYWTDSONKKEVISSDNQLQPLKQKSNRSKSTSGAPVPRDN 180
 QY 181 DGIPDSLEVEGYTVDPVKNKRTFLSPWISNIEHKGLTKYKSSPEKWSGASDPYDFEKT 240
 DB 181 DGIPDSLEVEGYTVDPVKNKRTFLSPWISNIEHKGLTKYKSSPEKWSGASDPYDFEKT 240
 QY 241 GRIDKNVSPERHPVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSTHT 300
 DB 241 GRIDKNVSPERHPVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSTHT 300
 QY 301 SEVHGNAEVHASFFDIGSVSAGFNSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
 DB 301 SEVHGNAEVHASFFDIGSVSAGFNSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
 QY 361 NANIRYVNTGAPINNVPTTSLVLGKQKQTLATDADENQISQILAPNNYFSPKNLAPIA 420
 DB 361 NANIRYVNTGAPINNVPTTSLVLGKQKQTLATDADENQISQILAPNNYFSPKNLAPIA 420

QY 421 LNAQKFFSPTITMYNQFLELEKTKQLRLDQVYGNIAFYNFENGRVVRVDTGSNWSEV 480
 DB 421 LNAQNDFFSPTITMYNQFLELEKTKQLRLDQVYGNIAFYNFENGRVVRVDTGSNWSEV 480
 QY 481 LPOIQETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGNPNGL 540
 DB 481 LPOIQETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGNPNGL 540
 QY 541 QYQKDKITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAMNLIIRDKEAFHYDR 600
 DB 541 QYQKDKITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAMNLIIRDKEAFHYDR 600
 QY 601 NNIAVGADESIVKEAHREVINSSTGLNLINIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 DB 601 NNIAVGADESIVKEAHREVINSSTGLNLINIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 QY 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVVYAVTKENTINPSENGDTSTNG 720
 DB 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVVYAVTKENTINPSENGDTSTNG 720
 QY 721 IKKILIFSCKGYEIG 735
 DB 721 IKKILIFSCKGYEIG 735

RESULT 9
 AAR60179
 ID AAR60179 standard; protein; 735 AA.
 XX AC AAR60179;
 XX DT 25-MAR-2003 (revised)
 XX DT 03-APR-1995 (first entry)
 XX DE Protective antigen of Bacillus anthracis.
 XX KW Anthrax; Bacillus anthracis; fusion protein; protective antigen;
 XX KW protective antigen; cell killing; targeting; targeting; pathogen;
 XX KW intracellular; HIV; human immunodeficiency virus; toxin.
 XX OS Bacillus anthracis.
 XX PN WO9418332-A2.
 XX PD 18-AUG-1994.
 XX PF 14-FEB-1994; 94WO-US001624.
 XX PR 12-FEB-1993; 93US-00021601.
 XX PR 25-JUN-1993; 93US-00082849.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Leppla SH, Klimpel K, Arora N, Singh Y, Nichols PJ;
 XX DR WPI; 1994-279753/34.
 XX DR N-PSDB; AAQ70180.
 XX PT Nucleic acid encoding anthrax toxin fusion protein - useful for
 PT targeting toxin to specific cells, eg for killing tumour cells or HIV-
 PT infected cells.
 XX PS Disclosure; Page 81-83; 124pp; English.
 XX CC The sequence encoding the protective antigen of Bacillus anthracis may be
 CC used in the construction of a nucleic acid which encodes a fusion protein
 CC comprising the anthrax protective antigen binding domain of the native
 CC anthrax lethal factor and a sequence encoding an activity inducing domain
 CC of a second protein. The fusion proteins are useful for the specific
 CC killing of tumour cells or the killing of cells infected with
 CC intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX

Db 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQTLATIKAKENQLSQLAPNYYPSKNLAPIA 420
 Qy 421 LNAQKFSSTPTTMNYQFLEKTKQLRLDQVYGNATYNGRVRVDTGNNSEV 480
 Db 421 LNAQDFSSTPTTMNYQFLEKTKQLRLDQVYGNATYNGRVRVDTGNNSEV 480
 Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIARFNEPNGNL 540
 Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIARFNEPNGNL 540
 Qy 541 QYQKDIETFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600
 Db 541 QYQKDIETFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600
 Qy 601 NNIAVGADES SVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 Db 601 NNIAVGADES SVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 Qy 661 DMLNSSLRODQKTFIDFKKYNDKPLIYSNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
 Db 661 DMLNSSLRODQKTFIDFKKYNDKPLIYSNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
 Qy 721 IKKILIFSCKGYEIG 735
 Db 721 IKKILIFSCKGYEIG 735

RESULT 7
 AAMS1487
 ID AAMS1487 standard; protein; 735 AA.
 AC AAMS1487;

XX 01-FEB-2002 (first entry)
 DE Anthrax PA mutant K397Q.
 XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
 XX Synthetic.
 OS Bacillus anthracis.
 XX Key Location/Qualifiers
 FH Misc-difference 397
 FT /note= "Wild-type Lys substituted by Gln"
 XX WC200182788-A2.
 XX 08-NOV-2001.
 XX 04-MAY-2001; 2001WO-US014372.
 XX 04-MAY-2000; 2000US-0201800P.
 XX (HARD) HARVARD COLLEGE.
 XX Collier RJ, Sellman BR;
 XX WPI; 2002-017725/02.
 XX
 PT Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.

Claim 4; Page; 77pp; English.
 CC The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections

CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAMS1483)
 XX
 SQ Sequence 735 AA;

Query Match 99.3%; Score 3749; DB 5; Length 735;
 Best Local Similarity 99.5%; Pred. No. 1.1e-239;
 Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESESSQGLLYFFDLNFOAPMWVTSTTGLDLSIPSSSELENIPSENOYF 60
 Db 1 EVKQENRLNESESSQGLLYFFDLNFOAPMWVTSTTGLDLSIPSSSELENIPSENOYF 60
 Qy 61 QSAIWGFIKVKKSDYTFATISADNHVTVMVDDQEVINKASNKRIRLSKGLYQIKIY 120
 Db 61 QSAIWGFIKVKKSDYTFATISADNHVTVMVDDQEVINKASNKRIRLSKGLYQIKIY 120
 Qy 121 QRENPTKGLDFKLYMTDSQNKKEVSSDNLOLPELKQKSSNSRKRKSTASGTVDPDRN 180
 Db 121 QRENPTKGLDFKLYMTDSQNKKEVSSDNLOLPELKQKSSNSRKRKSTASGTVDPDRN 180
 Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDPEKVT 240
 Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDPEKVT 240
 Qy 241 GRIDKNVSPARHPIVAAPIVHVDMENILSKNEDQSTCNTDSETRTISKNTSRTHT 300
 Db 241 GRIDKNVSPARHPIVAAPIVHVDMENILSKNEDQSTCNTDSETRTISKNTSRTHT 300
 Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTASL 360
 Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTASL 360
 Qy 361 NAMIRYVNTGTAPIYVNLPTTSLVGLKQTLATIKAKENQLSQLAPNYYPSKNLAPIA 420
 Db 361 NAMIRYVNTGTAPIYVNLPTTSLVGLKQTLATIKAKENQLSQLAPNYYPSKNLAPIA 420
 Qy 421 LNAQKFSSTPTTMNYQFLEKTKQLRLDQVYGNATYNGRVRVDTGNNSEV 480
 Db 421 LNAQDFSSTPTTMNYQFLEKTKQLRLDQVYGNATYNGRVRVDTGNNSEV 480
 Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIARFNEPNGNL 540
 Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIARFNEPNGNL 540
 Qy 541 QYQKDIETFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600
 Db 541 QYQKDIETFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600
 Qy 601 NNIAVGADES SVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 Db 601 NNIAVGADES SVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 Qy 661 DMLNSSLRODQKTFIDFKKYNDKPLIYSNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
 Db 661 DMLNSSLRODQKTFIDFKKYNDKPLIYSNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
 Qy 721 IKKILIFSCKGYEIG 735
 Db 721 IKKILIFSCKGYEIG 735

RESULT 8
 AAMS1489
 ID AAMS1489 standard; protein; 735 AA.
 XX
 AC AAMS1489;

XX 01-FEB-2002 (first entry)
 DE Anthrax PA mutant D425N.

CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAMS1483)
XX
SQ

Sequence 735 AA;

Query Match 99.4%; Score 3753; DB 5; Length 735;

Best Local Similarity 99.6%; Pred. No. 5.8e-240;

Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDYFATADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYFATADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRSTSGPTVPDRDN 180
Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKWSASDPYDFEKT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKWSASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDTPDQVYGNATYVNFNGRVRVDTGNSWSEV 480
Db 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDTPDQVYGNATYVNFNGRVRVDTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPGNL 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPGNL 540
Qy 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 600
Db 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 600
Qy 601 NNIAVGADESUVKEAHEVINSSTEGLLINIDKDIRKILSGYVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLINIDKDIRKILSGYVIEDETEGLKEVINDRY 660
Qy 661 DMLNISLQDQKTFIDFKKYNKLPYISNPNYKNVYAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISLQDQKTFIDFKKYNKLPYISNPNYKNVYAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 6

AAMS1490

ID AAMS1490 standard; protein; 735 AA.

XX AC AAMS1490;

XX AC AAMS1490;

DT 01-FEB-2002 (first entry)

XX Anthrax PA mutant D425E.
DE Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; muten.
XX Bacillus anthracis.
OS Synthetic.
XX Key Location/Qualifiers
FT Misc-difference 425 /note= "Wild-type Asp substituted by Glu"
XX WO200182788-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US014372.
XX 04-MAY-2000; 2000US-0201800P.
XX (HARD) HARVARD COLLEGE.
XX PA
XX PI Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
PS Claim 4; Page; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAMS1483)
XX
SQ Sequence 735 AA;

Query Match 99.4%; Score 3750; DB 5; Length 735;

Best Local Similarity 99.5%; Pred. No. 9.1e-240;

Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDYFATADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYFATADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRSTSGPTVPDRDN 180
Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKWSASDPYDFEKT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKWSASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
 |||||
 Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
 |||||
 QY 421 LNAQKDFSSPTITMNYNOFLELEKTKQLRLDQVYGNIAIYNPENGVRVDTGNNWSEV 480
 |||||
 Db 421 LNAQKDFSSPTITMNYNOFLELEKTKQLRLDQVYGNIAIYNPENGVRVDTGNNWSEV 480
 |||||
 QY 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAFGFNEPKNL 540
 |||||
 Db 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAFGFNEPKNL 540
 |||||
 QY 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAKKNILIRDKRFHYDR 600
 |||||
 Db 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAKKNILIRDKRFHYDR 600
 |||||
 QY 601 NNIAVGADESUVKEAHEVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
 |||||
 Db 601 NNIAVGADESUVKEAHEVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
 |||||
 QY 661 DMLNISSLRQDGKTFIDPKKYNKDLPLXISPNYKNVYAVTKENTIINPSENGDTSTNG 720
 |||||
 Db 661 DMLNISSLRQDGKTFIDPKKYNKDLPLXISPNYKNVYAVTKENTIINPSENGDTSTNG 720
 |||||
 QY 721 IKKILIFSCKGYEIG 735
 |||||
 Db 721 IKKILIFSCKGYEIG 735
 |||||

RESULT 3

AAM51485
 ID AAM51485 standard; protein; 735 AA.

XX AC AAM51485;
 DT 01-FEB-2002 (first entry)

XX Anthrax PA mutant K397D.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; muten.

XX Bacillus anthracis.

XX Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 397 /note= "Wild-type Lys substituted by Asp"

XX WC200182788-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201900P.

XX (HARD) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

XX Protecting humans against anthrax using mutant B groups (anthrax

XX protective antigens) of the pore-forming binary A-B toxin of Bacillus

XX anthracis.

XX Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of

XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants

XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B

XX

CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections
 CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)

XX SQ Sequence 735 AA;

Query Match 99.5%; Score 3755; DB 5; Length 735;

Best Local Similarity 99.6%; Pred. No. 4.2e-240;

Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLLNESSSQGLLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSNQYF 60

Db 1 EVKQENRLLNESSSQGLLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSNQYF 60

QY 61 QSAIWGFIKVKSDSYTATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKQY 120

Db 61 QSAIWGFIKVKSDSYTATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKQY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRKSTAGTVPDRDN 180

Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRKSTAGTVPDRDN 180

QY 181 DGIPDSLEVEGYVDVKNKRTFLSPMISNHEKGLTKVKSSEKASTADPYSDPEKVT 240

Db 181 DGIPDSLEVEGYVDVKNKRTFLSPMISNHEKGLTKVKSSEKASTADPYSDPEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEQDQSTNDSETRTTSKNTSTSRHT 300

Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEQDQSTNDSETRTTSKNTSTSRHT 300

QY 301 SEVHGNAEVHASFFDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Db 301 SEVHGNAEVHASFFDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420

Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITMNYNOFLELEKTKQLRLDQVYGNIAIYNPENGVRVDTGNNWSEV 480

Db 421 LNAQKDFSSPTITMNYNOFLELEKTKQLRLDQVYGNIAIYNPENGVRVDTGNNWSEV 480

QY 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAFGFNEPKNL 540

Db 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAFGFNEPKNL 540

QY 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAKKNILIRDKRFHYDR 600

Db 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAKKNILIRDKRFHYDR 600

QY 601 NNIAVGADESUVKEAHEVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660

Db 601 NNIAVGADESUVKEAHEVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660

QY 661 DMLNISSLRQDGKTFIDPKKYNKDLPLXISPNYKNVYAVTKENTIINPSENGDTSTNG 720

Db 661 DMLNISSLRQDGKTFIDPKKYNKDLPLXISPNYKNVYAVTKENTIINPSENGDTSTNG 720

QY 721 IKKILIFSCKGYEIG 735

Db 721 IKKILIFSCKGYEIG 735

RESULT 4

AAM51491

ID AAM51491 standard; protein; 735 AA.

XX AC AAM51491;

XX DT 01-FEB-2002 (first entry)

XX

CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
SQ Sequence 735 AA;

Query Match 100.0%; Score 3774; DB 5; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.3e-241; Indels 0; Gaps 0;
Matches 735; Conservative 0; Mismatches 0;

QY 1 EVKQENLLNESSESSQGLLGYYFSDLNFAQPMVVTSSITGDLSPSELENIPSENQYF 60
Db 1 EVKQENLLNESSESSQGLLGYYFSDLNFAQPMVVTSSITGDLSPSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPLKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPLKQKSSNSRKRSTASGTPVDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKSTASDPYDFEKT 240

QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEQDSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEQDSTQNTDSETRISKNSTSRHT 300

QY 301 SEVHGNAEVHAFDIDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAFDIDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRVNTGTAPIYVNLSTSLVGLKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYVNLSTSLVGLKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKFSSTPIITMNYNQFLELEKTQKLELDTQVYGNIAATYFNFGRVVRVDTGNSWSEV 480
Db 421 LNAQKFSSTPIITMNYNQFLELEKTQKLELDTQVYGNIAATYFNFGRVVRVDTGNSWSEV 480

QY 481 LPOIQETTARIIFNGKDLNVERRTAAVNPSPLETTKPDMTLKEALKIATGFENPKNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRTAAVNPSPLETTKPDMTLKEALKIATGFENPKNL 540

QY 541 QYQKIDITEFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKLNAKMNIILRDKRFHYDR 600
Db 541 QYQKIDITEFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKLNAKMNIILRDKRFHYDR 600

QY 601 NNTAVGADESVKBAHREVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNTAVGADESVKBAHREVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660

QY 661 DMLNISLRDQGTFFIDFKYNDKLPYISNPNYKVVYAVTKENTIIINPSENGTSTING 720
Db 661 DMLNISLRDQGTFFIDFKYNDKLPYISNPNYKVVYAVTKENTIIINPSENGTSTING 720

QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 2
AAM51493
ID AAM51493 standard; protein; 735 AA.
XX
AC
XX AAM51493;
XX
DT 01-FEB-2002 (first entry)

XX Anthrax PA mutant K397D/D425K.
DE
XX Anthrax; PA: protective antigen; antibacterial; pore-forming toxin;
KW moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.
XX
OS Bacillus anthracis.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"
FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"
XX
XX WO200182788-A2.
XX
XX 08-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014372.
XX
XX 04-MAY-2000; 2000US-0201800P.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
PT anthracis.
XX
XX Claim 4; Page: 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
SQ Sequence 735 AA;

Query Match 99.7%; Score 3761; DB 5; Length 735;
Best Local Similarity 99.7%; Pred. No. 1.7e-240; Indels 0; Gaps 0;
Matches 733; Conservative 0; Mismatches 2;

QY 1 EVKQENLLNESSESSQGLLGYYFSDLNFAQPMVVTSSITGDLSPSELENIPSENQYF 60
Db 1 EVKQENLLNESSESSQGLLGYYFSDLNFAQPMVVTSSITGDLSPSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPLKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPLKQKSSNSRKRSTASGTPVDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKSTASDPYDFEKT 240

QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEQDSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEQDSTQNTDSETRISKNSTSRHT 300

QY 301 SEVHGNAEVHAFDIDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAFDIDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Tue May 4 14:31:02 2004

Db 1395 NKNHNFINSYFDLNENEKKKINIF 1419

Search completed: May 3, 2004, 19:39:59
Job time : 37.6261 secs


```
QY 682 NDKPLVYINPNKVNVAVTKNTIINPSENGDTS 717
DB 4416 SNKLP-CLNNINYQIK---INKSHYII--SKNGEWS 4445

RESULT 13
Q8IKS2 ID Q8IKS2 PRELIMINARY; PRT; 1904 AA.
AC Q8IKS2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ferlin, putative.
GN PF14_0530.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrall B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum";
RL Nature 419:498-511(2002).
EMBL: AE014825; AA37143.1; -.
DR InterPro; IPR000008; C2.
DR Pfam; PF00168; C2; 5.
DR SMART; SM00239; C2; 5.
DR PROSITE; P85004; C2 DOMAIN 2; 2.
SQ SEQUENCE 1904 AA; 224720 MW; AA36470367BDEC41 CRC64;

Query Match 5.4%; Score 203.5; DB 5; Length 1904;
Best Local Similarity 19.8%; Pred. No. 0.28;
Matches 174; Conservative 129; Mismatches 327; Indels 247; Gaps 43;

QY 3 KOENRLNSESSESSQGLGYPSDLN-FQAPMVVTSSTTGDLSIPSSSELENIPSENQVFO 61
DB 836 KDDNN--NKKDNNNNNNNNYNSNVYQNDLLYGERVGDILL---YFELVQSKDA-MK 889

QY 62 SAIWSGFIKVKSDSEYTFATSDADNHVTVWDDQEVINKASNKIRLEKGRLYQIK---- 117
DB 890 FPIYPMITEIKKCTLSFFQMSLENILM-----KXANFLKTLSPFRNNKYQISTPII 941

QY 118 -----IQORENPTEKGLDPKLYWTDSONKK-EVISSDNLQLPELK 157
DB 942 LLSITSYSYSGKKKELMIKYEKTLKANTRIQLK-NWKNSTNQSQSFENFSIENWNI----- 996

QY 158 QKSSNSRKRKSTSGAGTVPDRNDGIPDS-LEVEGYTVDVKNK-----RTFLS-----PWI 207
DB 997 -----DPLDPFDPILNIKVKYKSKYFGETNLSLVPYLPWI 1037

QY 208 SNIHEKKGLTKYKSSPEKSTASDPYDFEKV-----TGRIDKNVSPEARHPLVAAPYIV 262
DB 1038 KNIDE-----VLV-----YLQHDYSETINKNDTNTFYKN-----KNAALVISALS 1082

QY 263 HYDMENILSKNEDOSTQNTDSET-----RTISKNTSTRTHTSEVHGNAEVHASPFD 315
DB 1083 LADCEDTSLSEKINEKYENDDDEAWKEIPLFLNDQENQEDKNKNTSSQHGNTVNNYDGYN 1142

QY 316 IGSYSVAGFSNSNSTVAIDHLSLAGRTWAETWGLTACTARLANAIRVYNTGTAPI- 374
DB 1143 -NGAYEMGMNYMETNINKNDN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1192

QY 375 ----YNNVPTTSLVLGKNTLATIKAD--ENQLSQILAPNNYPPSKGLAPIALNAQKOPSS 429

Db 1193 SYNNVL-----QNDTNNVRYNHSNNMMNNNNYNNASQGVNNYNNYND 1244
QY 430 TPITMAYN-----QFLELEKTK-----QLRLTDDQVYGNATYN-FENGREVAVDTGSNW 477
Db 1245 KGNLNFNNNNHHFNKLSNKKFSDSYLSRQKDTYNIKNNNSYKLFDDGIPPIIKLSYN 1304
QY 478 SEVLPOIQETTARIIFN-----GKDLNLVERRIAAYN----- 509
Db 1305 VKNVPYKILTSKYLNVHIPPFRILYVEGDKLN-TEKFKINRVSDVGLILENLYDDIL 1363
QY 510 -PSDPLETTKPDMTL-----KEALKIA-----FGNEPNNG-----NLQYQOKDI 547
Db 1364 IPSLPLIKKNCISDNNYNNENKIEKGKFGCFEQFPFVEIIGGQIKCFKTKIKYNLES 1423
QY 548 TEFDFNPDQOTSONI-----KNQLA-----ELNATNIYTVLDKIKNAKMILIR 592
Db 1424 ENMPLSLKDTNQNIIFRNKFRGKNIPLYLKIRVYVLRGILGYGINNEYTANPYLIFSLG 1483
QY 593 DK-----RFHYDRNNI-----AVGADSVVKEAHREVINSTEGLLNLNDKIRKILSGYI 643
Db 1484 EKTSLNRAFGRSNINPEFQCLWESAIFFE--DEILTISVSAEDNYDKINDIYIG-- 1539
QY 644 VEIETDEGLKEVIN--DRYDMLNLISSLRDQKTFIDPKK-YN--KLPLVISNPNKV-- 696
Db 1540 ----SIE-----INLFDRWMSKEWRHMKCKIPVEYRPLYSNYIKHPKXVSSNNYNTWN 1590
QY 697 ---NVYAV-----TKENIINPSENGDTS 717
Db 1591 SMNNIFSFDFIYNYLTYSPTKGNNNNNNNNNNS 1627

RESULT 14
Q8IHY4 ID Q8IHY4 PRELIMINARY; PRT; 2849 AA.
AC Q8IHY4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pf11_0392.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrall B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum";
RL Nature 419:498-511(2002).
EMBL: AE014841; AA35975.1; -.
KW Hypothetical protein.
SQ SEQUENCE 2849 AA; 335916 MW; B5515D173D96B813 CRC64;

Query Match 5.3%; Score 200.5; DB 5; Length 2849;
Best Local Similarity 20.6%; Pred. No. 0.65;
Matches 170; Conservative 133; Mismatches 334; Indels 189; Gaps 41;

QY 3 KOENRLNSESSESSQGLGY-IFSDLNFO-APMVVTSSTTGDLSIP--SSELENIPSEN 57
Db 1862 KHNYTYFNKEENLKSIFKYNNNNNINNEKDDIPKSIQSSFINEDNATYNNKGNMKJK 1921
QY 58 Q-YFQSAIWSGFIKVKKSDSEYTFATSDADNHVTVWDD-----QEVINKASNSNK---- 105
```



```
QY 226 WSTASDPYDFKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEQDSTQNTDSE 285
DB 231 PCTANDPYDFKVSQIDPSVSMVARDPMISAYPIGVQMERLVVSKSE-----TITGDS 286
QY 286 TTTISQNTSTSTHSEVHGNAEVAHSPFDIGG-----SVSAGFSNSNSTVAIDHSL 338
DB 287 TKSMKSSTSHSNTNTV--GAEVGSLQLAGIIPVFSMSASANSYHTQNTSVDDT- 343
QY 339 SLAGERTWAETMGLNTADTARLANAIRYNTGTAPIYVLTPTTSLVLGKQNTLATIKAE 398
DB 344 --TGE-SFSQGLSINTGESAYINPNIYNTGTAPVYNTPTTTIVIDK-QSVATIKQGE 399
QY 399 NOLSQLAPNNYVPSKNLAPIALNAQKOPSSPTITMNYNQFLELEKTKQLRLDQVYGN 458
DB 400 SLIGDYLNGGTYPIIGEPFPMALNTWDQFSSRLIPINYNQLKSIDNGGTVMLSSTQFTGN 459
QY 459 IATYFENGRRVVDGTGSNWSEVLPOIQTETARI--FNGKOLNVERIAAVNPSDPLET 516
DB 460 FAKYN-SGNLVTD-GNNWGPYLGITKSTASLTLSFGQTTQVA--VWAPNPSDPEX 514
QY 517 TKPDMTLKALKIATFG--FNEPNGNLOVQGDIT--BEDPFPQQTQSONIKQLAELNATN 573
DB 515 T-PKLTLEQLVKAFALEKNGKFFPHGLEISKNEKIQVFLDSNTNDFENQKNTAKD 573
QY 574 IYVLDKIKLNAQONILIRDKRPHYDRNNIAV 605
DB 574 IMHCI--IKRN--MNLVVKVITFKENISSINI 601

RESULT 8
ID Q844J8 PRELIMINARY; PRT; 775 AA.
AC Q844J8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Viplac.
GN Viplac.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y., Chen J., Pang Y.;
RT "Cloning of vipA(C) and vip2A(C) from Bacillus thuringiensis HD201.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY245547; A086514.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxins.
DR Pfam; PF03495; Binary toxin.
DR PRINTS; PR01391; BINARYTOXIN.
DR SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;

Query Match 19.3%; Score 728; DB 2; Length 775;
Best Local Similarity 29.1%; Pred. No. 1.3e-26;
Matches 224; Conservative 133; Mismatches 293; Indels 120; Gaps 30;

QY 4 QENRLNESSSGGLIGYFSDINFPQPMVWTSSTGDLSPSSLEN--IPSENQYFQ 61
DB 42 QKQKQ--QKEMDRKGLLGYFFKQDF-SNLTMPFAPTRNTLIYDQQTANKLLDKKQQYQ 97
QY 62 SATWSGFIKKKSDDEVTFATSDAHNVTWYDDDEVINKASNSKIRLEKGRLYQIKIQYQ 121
DB 98 SIRWIGLIQKKGDFTFNLSEBQALIEIDGKIISNKGKQVHLEKLVPIKIEYQ 157
QY 122 RENPTEKGLD-----PKLVWTSQNKKEVISSDNLQLPELKQKSN----- 162
DB 158 SD--TKFNIDSKTFKPEKFKIDSQNSQOVQKRDENLPEFNPKESREFLAKASKTNPFM 215
QY 163 SRKKSSTAGPTVDRDNDGIPDSLEVEGVTVDVKNKRTLSLPMWNIHEKGLTKYKSS 222
```

```
DB 216 QKMKRDIDED---TDTGDSIPDLWEENGYT--IQNK--VAVKW-DDKFAQQGVYKYLSS 267
QY 223 PEKWSSTASDPYDFKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEQDSTQNT 282
DB 266 PYQARTVGPYTDWEKAAGDIPKSNAAATRNPLVAAPPSINVMRCKILLKSDSNLS---- 323
QY 283 DSETRTISQNTSTSTHSEVHGNAEVAHSPFDIGGSVAGFSNSNSS---TVAIDHSL 339
DB 324 -----NSAEAHNSNSYTYANSE-----GASIEAGFGPKGFSFGVSANTQHTET 366
QY 340 LAGERTW-----ABTMGLNTADTARLANAIRYNTGTAPIYVLTPTTSLVLGKQNTLATIK 395
DB 367 VGSQ--WGNSSKSNTEQFNSASAGYLNANVHYNNVGGIYDAQPTTSFIL-QDSTIATIT 423
QY 396 ADENQLSQLAPNNYVPSKNLAPIALNAQKDFSSPTITMNYNQFLELEKTKQLRLDQV 455
DB 424 AKSNATALSFPSGDRYPASK-EGISLKTWDDFNSHPITLTKPQLDVAVLANNEVIKINTDQT 482
QY 456 YGNATYFENGRRVVDGTGSNWSEVLPOIQTETARIIFNGKOLNVERIAAVNPSDPLE 515
DB 483 DGRYGIIGVDG--KAEIGDRNSPIIDEIKGTASIIIDPADGKALETRIAAKDYKNPD 539
QY 516 TTKPDMTLKALKIATFG--FNEPNGNLOVQGD---ITEFDNF-----FDQQTSON 561
DB 540 KT-PSLTIKEGLIKAYPESISEDKGLFVEYKNDGKVTKKQLSEENIMPYLDEDTKE 598
QY 562 IKQLAELNATNIYVLDKIKLNAQONILIRDKRPHYDRNNIAVGADESIVVKEAHEVIN 621
DB 599 FERQLSDGSAGLY---DIKLTPEKNITIR-----LATVTLGFDQDFSAYPENATW 647
QY 622 SSTEGLL-----LNIDKDIRK-----ILSGYIVE--IEDTEGLKEVI---NDR 659
DB 648 SDRFNGNLRLGSLAIPQESKTYIPDKVKPNYDYLITGYIKHFTTDSLSLGIVAFKKDN 707
QY 660 YDMLNTSSLRQDGKTFIDFKKYNDKLPYISNPKNYKVVAVYKENTIIIN 709
DB 708 FEMXNMGTSIFPSONSGEGFKFTIKQ-NISG-DYILDSIQLMKRNDVN 755

RESULT 9
ID Q8KYK2 PRELIMINARY; PRT; 225 AA.
AC Q8KYK2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protective antigen-related protein, (px01-111).
GN BXA0163.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2012;
RX MEDLINE=22061436; PubMed=12004073;
RA Read T.D.; Salzberg S.L.; Pop M.; Shumway M.; Umayam L.; Jiang L.;
RA Holtzapfel E.; Busch J.D.; Smith K.L.; Schupp J.M.; Solomon D.;
RA Keim P.; Fraser C.M.;
RT "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in
RT Bacillus anthracis.";
RL Science 296:2028-2033(2002).
DR EMBL; AB011190; AAM26108.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 225 AA; 25402 MW; 2E121BE54295F9C8 CRC64;

Query Match 7.9%; Score 296.5; DB 2; Length 225;
Best Local Similarity 37.2%; Pred. No. 6.9e-07;
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;

QY 566 LAELNATNIYVLDKIKLNAQONILIRDKRPHYDRNNIAVGADESIVVKEAHEVINSTE 625
```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 10748;
RX MEDLINE=94041537; PubMed=8225592;
RA Perelle S., Gibert M., Boquet P., Popoff M.R.;
RT "Characterization of Clostridium perfringens iota toxin genes and
RL expression in Escherichia coli.";
RN Infect. Immun. 61:5147-5156(1993).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 10748;
RA Popoff M.R.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X73562; CAA51960.1; -.
DR PIR; I40862; I40862.
DR HSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0003405; P:pathogenesis; IEA.
DR InterPro; IPR003896; P:anthrax toxinB.
DR Pfam; PF03495; B:binary_toxB; I.
DR PRINTS; PR01391; BINARYTOXINB.
KW SIGNAL.
FT SIGNAL 34 38 POTENTIAL.
FT CHAIN 212 875 IOTA TOXIN COMPONENT IB.
SQ SEQUENCE 875 AA; 98468 MW; C9AE092CD3618921 CRC64;

Query Match 23.2%; Score 876; DB 2; Length 875;
Best Local Similarity 31.2%; Pred. No. 1.5e-33;
Matches 253; Conservative 130; Mismatches 278; Indels 150; Gaps 31;

QY 1 EVKQENRLNBSSESSQGLGYFSDLNFOAPMVVTSSTGSLSPSELENIPSE-NOY 59
DB 35 DTNQKEIINTNLNSGLMGYFADHPKDELPAPKNGDLKFEKKVLLTEDNSS 94

QY 60 FQSAIWGPIKVKSDYEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKI 119
DB 95 IKSIRWTGRIIPSEDGEVILSTDR-NDVLNQINAKGDIK---TLKVNMGKQAYNIRIE 150

QY 120 YQREN-----PTEKGLDEKLYWTDSONKKEVSSDNLQLPKQKSSNRKRSTS 170
DB 151 IQDKNGSIDNLSVP-----KLYW-ELNGNKTVIPENLFFROYKIDEND----- 195

QY 171 AGPTVP-----DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIH 211
DB 196 --PFIPINNFFVRPFSAAWEDELDTDNDNIPDAVEKNGYTI---KDSIAVKWDSFA 249

QY 212 EKGLTKYSSPKWSTADPYSDPEKVTGRDKQVSPPEARPLVAAYPIVHVDNENIL 271
DB 250 E-QGYKKYSSYLESTAGDPYTDYQKASGSDKAIKLEARDPLVAAYPVVGVGMENLII 308

QY 272 SKNEDOSTQNTDSERTIKNTSTSTSTSEVHGNAEVAHSPFDIGGSVSAFSSNSST 331
DB 309 STNEHASS-----DQKTVSEATNTSKTDANTV-----GVSIAGYQNGFTGN 351

QY 332 VAIDHS-----LSAGERTWAETMGNTADTARLNANIRYVNTGTAPIVNVLTSL 383
DB 352 ITTSYHTTNDSTAVQDSNGESWNTGLSINKGESAYINANRYNTGTAPMYKVPTTNL 411

QY 384 VLGNQOTLATIKADENQLSQILAPNNYPSKNLAPALNAQKDFSTPTIMYNOFLELE 443
DB 412 VL-DGETLATIKADQNGIGNLSNPNETPKGLSPALNTMQFNARLLPINYDQKKLD 470

QY 444 KTKQLRLDTPQVYGNATYVNGRVRVDTGSNWSEVLPCQIBETTARIIFNGKDLNVER 503
DB 471 SGKOIKLETTQVSGNYGTKN-SQGI-IIEGNSWNYISQIDSVSASIILD-TGSQTFER 527

QY 504 RIAVNPSPDLETKPDMLKEALKAFNGENGNLQV-QGQDITE--PDFNFDQTSQ 560
DB 528 RVAKEGNGEDKT-PEITIGEAIKAFSATK-NGELLYFNGIPIDESCVELIFDNTSE 585

QY 561 NIKNQLAELNATNIYVLDIKLNKAKNLIIRDKRF--HYDR-NNIAVGADESVMKEAHR 617
DB 561 NIKNQLAELNATNIYVLDIKLNKAKNLIIRDKRF--HYDR-NNIAVGADESVMKEAHR 617

```

```

DB 586 IIEKQKYLDDKKIYV-----KLERGMNLIKVPYSYFTNFDYNNFP--ASMSNIDTKNQ 639
QY 618 EVINSTEGL-----LLNIDKDIRKILSGY-----IVEIEDTGLKEVINND 658
DB 640 DGLQSVANKLSGETKIIIPMSKLPKYKVPFGSKDPSTNSITVNIKSKEQKTDYLPV 699
QY 659 RYDMLNIS-----SLRQDKTIDFKYND--KLPLYISNPYKV----- 696
DB 700 EKDYTKFSYEFETTGKSSDIEITLSSGVIFLDNLSITELASTPEILKEPIKVPDQOE 759
QY 697 -----NVYAVTKENTIINPSENGDTSTNGI 721
DB 760 ILDAHNNKYADIKLDT-----NTGNTYIDGI 785

RESULT 7
ID O86171 PRELIMINARY; PRT; 721 AA.
AC O86171;
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE C2 toxin (Component-II).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=36323874; PubMed=9659689;
RA Kimura K., Kubota T., Ohishi I., Isogai H., Isogai E., Fujii N.;
RT "The gene for component-II of botulinum C2 toxin.";
RL Vet. Microbiol. 62:27-34(1998).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=36184657; PubMed=8645309;
RA Fujii N., Kubota T., Shirakawa S., Kimura K., Ohishi I., Moriishi K.,
RA Isogai E., Isogai H.;
RT "Characterization of component-I gene of botulinum C2 toxin and PCR
RT detection of its gene in clostridial species.";
RL Biochem. Biophys. Res. Commun. 220:353-359(1996).
DR EMBL; D89982; BAA32537.1; -.
DR HSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0003405; P:pathogenesis; IEA.
DR InterPro; IPR003896; B:anthrax toxinB.
DR Pfam; PF03495; B:binary_toxB; I.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 721 AA; 80515 MW; 44C8153AC749D5F2 CRC64;

Query Match 21.7%; Score 818; DB 2; Length 721;
Best Local Similarity 35.4%; Pred. No. 6.4e-31;
Matches 224; Conservative 117; Mismatches 227; Indels 64; Gaps 25;

QY 3 KOENRLNBSSESSQGLGYFSDLNFOAPMVVTSSTGSLSPSELENIPSENOYF 60
DB 5 KFENVSKNSKNYPTINGLAGYFEN-DFFNLNIIISPTDGNLTFSKEDINSILG-NKII 62

QY 61 QSAIWSGFIKVKSDYEYTFATSDN-HVTMWVDDQEVINKASN-SNKIRLEKGRLYQIKI 118
DB 63 KSAEWIGLIPKSIITGEVILSTNSPNCVEL--NGEIFNLSLNTSNTVNLIQNVYDRI 119

QY 119 -QYORENTEKGLD-FKLYWTDSONKKEVSSDNLQLPKQKSSNRKRKSTSGPTV- 175
DB 120 EQLMSENQLLNKYGKIKLYMETSDIIKEIIPSEVL---LKNPNYSNTNKSKEKFPNNTLF 175

QY 176 -----PDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEK 225
DB 176 SNAKLANANRDTDRDGIPEDEWEINGITVMNQKAVAMDDKFAAN-----GYKTVSNPFK 230

```

Db 321 ----DQKTVSRNTNSTKDANTAGVAINIAYQNGF-----TGSITTNVSHHTENSTAVONS 373
 QY 338 LSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLVLGKNTQTATIKAD 397
 Db 374 ---NGB-SWNTSLKINGSAYINAVRYNTGTAPMYKVTPTNLVL-DGDTLTITKAQ 428
 QY 398 ENQLSQILAPNYYSKLAPALNAQKDFSSPTITMNYNQFLEKTKQLRLDTDOYVG 457
 Db 429 DNOIGNNLSPNRTYPKKGLSPLALNTMOFSSRLPIYNDQKLDAGKQIKLETTQVSG 488
 QY 458 NIATYNFENGRVRVT-GSNWSEVLPOQIETTARIIFN-GKDLNLVERRIAAVNPSPLE 515
 Db 489 N---YGIKNSQOIIITEGNSWSDYISQIDSLASIIIDTGS--VFERRVTAKDSSNPED 543
 QY 516 TTKPDMTLKEALKIAPGNEPNGNLOYQKDKITE--FDNFDOQTSQNIKNQLAELNATN 573
 Db 544 KT-PVLTIGEATEKAFGATKNGEILYFNGMPIDESCVELIFDGNTANLIKERNALNDKK 602
 QY 574 IYTVLDKIKLNAKMLIIRD-----FSGENVMYKRLIYAITPEN-----KELLV 875
 Db 603 IYVNV---QLERGKMLIKTSTYFNNFDGYNFPSSWNSVDSNNDGQLQNAANKLSGETK 658
 QY 594 -----KRPY-----KRPY----- 598
 Db 659 IVIPMSKLNPKRYVPFSGYLKNSSTNPITVNIKABQKTYNLVSENDYKFPSEYEFITG 718
 QY 599 -DRNN-----IAGADESVKKEAUREV-----I 620
 Db 719 RDSNIEITLSSGTIFDLNLSITELNSTPEILKPDIKVPDQELI-DAHKYVADLSF 777
 QY 621 NSSTEGLLN-----IDKDIRKILSGYIVIE-DTEGLKEVINDRYDMLNLSRLQD 671
 Db 778 NQSTANYLDGLYFEPTQNTKEVLDYIQYKVEATLEYSGFDIGTKDKELANYTGDSNQ 837
 QY 672 GKT-FIDEKYNKDLPLVYISNPNY-----KVVVAVTKENTIINPSENGTSTNGIKKILI 726
 Db 838 PKTYNVNFRSY-----FSGENVMYKRLIYAITPEN-----KELLV 875
 QY 727 FS 728
 Db 876 LS 877

RESULT 5
 Q9KH41 ID Q9KH41 PRELIMINARY; PRT; 876 AA.
 AC Q9KH41;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE CdcB.
 GN CDTB.
 OS Clostridium difficile.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 CX NCBI_TaxID=1496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CCUG 20309;
 RA Chang S.Y., Song K.P.;
 RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain
 CCUG 20309.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF271719; AAF81761.1; -.
 DR HSSP; P13423; 1ACC.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0015070; F:toxin activity; IEA.
 DR GO; GO:0009405; P:pathogenesis; IEA.
 DR InterPro; IPR003896; Anthrax toxinB.
 DR Pfam; PF03495; Binary toxinB; I.
 DR PRINTS; PR01391; BINARYTOXINB.
 SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;

Query Match 23.3%; Score 880.5; DB 2; Length 876;
 Best Local Similarity 31.6%; Pred. No. 9e-34;
 Matches 251; Conservative 139; Mismatches 280; Indels 125; Gaps 30;
 QY 4 QENRLNSESSESSQGLGYFSDLNFOAPMVYTSSTGDSIPSSBELEN-IPSENOYFQS 62
 Db 40 KKKEIVNEIDILPNGLMGYYFTDEHFQDLKMAPIKDGNLKEFEKKVLDKDKSDVKS 99
 QY 63 AHWGIGIKVKKDEYTFATSAADNHVTMWYDDQEVINKASNSNKIRLEKRLVQIKIYOR 122
 Db 100 IRWTGILIPSKGEGYFLSDRDD-VLMQVNTSTST---SNTLKVNNKGEKVKVRIEIQD 155
 QY 123 EN--PTEKGLDFKLYWTDQSNKKEVISSDNLOLPELKOKSSNSRKRSTSACTVP---- 176
 Db 156 KNLGSDNLSSPLXW-ELDGMKKIIPENLFLRDY-----SNIKD-----DPFIPNNNF 205
 QY 177 -----DRDNGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKKS 222
 Db 206 FDPKLSMDWEDELDTDNDNIPDSYBERNGYTI-----KLIIVKWDSEFAE-QYKKYVSN 250
 QY 223 PEKWSSTASDPYSDFEKTGRIDKNSVPEARHPLVAAPIVHYVDMENIILSKHEDQSTONT 282
 Db 261 YLESNTAGDPYTDYKASGSFDKAIKTEARDPLVAAPIVGVGMEKLIISTNEHAST--- 317
 QY 283 DSETRISKNTSTSRTHTSEVHGNAEVHASFPDIGSVSAGFSNSNSTVAIDHS----- 337
 Db 318 -DQKTVSRATNTSKTESNTA-----GVSVMVYQNGTANVTNYSTTTON 363
 QY 338 ---LSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLVLGKNTQTATIK 394
 Db 364 STAVDSNGSNGWNTG-SINKGESAYINAVRYNTGTAPMYKVTPTNLVL-DGDTLTIT 422
 QY 395 KADENQLSQILAPNYYSKLAPALNAQKDFSSPTITMNYNQFLEKTKQLRLDTDO 454
 Db 423 KAQENQIGNNLSPGDTPYKGLSPLALNTMOFSSRLPIYNDQKLDAGKQIKLETTQ 482
 QY 455 VYGNIAIYFENGRVRYDTGNSWSEVLPOIETTARIIFNGKDLNLVERRIAAVNPSPDPL 514
 Db 483 VSGNFGTKN-SSGOI-VTEGNSWSDYISQIDSLASIIIDLTEN-ESYERRVTAKNLODPE 539
 QY 515 ETTKPDMTLKEALKIAPGNEPNGNLOYQKDKITE--FDNFDOQTSQNIKNQLAELNAT 572
 Db 540 DKT-PELTIGEATEKAFGATKDGILYFNDIPIDESCVELIFDDNTANKIKSLKLSDK 598
 QY 573 NIYTVLDKIKLNAKMLIIRDKRPY---DRNIIAVGADESVMVKEAUREVINGSTEG-- 627
 Db 599 KIYVNV---KLERGMNLIKTPYFTFNDDYNNVP--STWSNVNTNKGDLQGSANKLNG 652
 QY 628 -----LLMIDKDIRKILSGY-----IVEI-----EDTEGL-----KEVI 656
 Db 653 ETIKIKIPMSSELKPKRYVPFSGYSKDPLTNSIIVKIKAKEEKTDLVPEQGYTKFSYEFE 712
 QY 657 NDRYDMLNIS-SLRQDKTFTIDFKKYN--KLPLVYISNPNYK-----VNVVAVTKENT 706
 Db 713 TTEKSSNIEITLIGSGTITLDNLISITELNSTPEILDEPEVKIPTDQEIIDAHKIYFADL 772
 QY 707 IINPSENGDTSTNGI 721
 Db 773 NFNPS-IGNTYINGM 786
 RESULT 6
 Q46221 ID Q46221 PRELIMINARY; PRT; 875 AA.
 AC Q46221;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Iota toxin component Ib precursor.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 CX NCBI_TaxID=1502;

Db 20 EVKQENRLNSESSESSQGLGYFSDLNQFAPMVVTSSTTGDLSIPSELENIPSENYQF 79
Qy 61 QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 80 QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 139
Qy 121 QRENPTTEKGLDFKLYWTDSONKKEVISSDNQLQPELKQSSNSRKKSTASGPTVPDRN 180
Db 140 QRENPTTEKGLDFKLYWTDSONKKEVISSDNQLQPELKQSSNSRKKSTASGPTVPDRN 199
Qy 181 DGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSSPEKWSSTASDPSDFEKT 240
Db 200 DGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSSPEKWSSTASDPSDFEKT 259
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Db 260 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 319
Qy 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 320 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 379
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420
Db 380 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILLAPNNYPSKNLAPIA 439
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
Db 440 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 499
Qy 481 LPQIQTETARIIIFNGKDLNVERRIAANVPSDPLETTPDWTLEKALKIAFGFNEPKNL 540
Db 500 LPQIQTETARIIIFNGKDLNVERRIAANVPSDPLETTPDWTLEKALKIAFGFNEPKNL 559
Qy 541 QYQKDIETEFDFNFDOQTSONIKQNLAEALNATYITVLDKIKLNAMNILLIRDKRFHYDR 600
Db 560 QYQKDIETEFDFNFDOQTSONIKQNLAEALNATYITVLDKIKLNAMNILLIRDKRFHYDR 619
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLINIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 620 NNIAGADESVVKEAHEVINSSTEGLLINIDKIRKILSGYIVEIEDTEGLKEVINDRY 679
Qy 661 DMLNLSLRDQGTFFIDFKYNDKPLYSNPNYKVVVAVTKENTIIINPSENGDTSTNG 720
Db 680 DMLNLSLRDQGTFFIDFKYNDKPLYSNPNYKVVVAVTKENTIIINPSENGDTSTNG 739
Qy 721 IK 722
Db 740 IK 741

RESULT 2
Q937W3 ID Q937W3 PRELIMINARY; PRT; 743 AA.
AC Q937W3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pag protein (Fragment).
GN PAC.
OS Bacillus anthracis.
OC Plasmid pX01.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Carbosap.
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
RA Fasanella A., Francia M., Ciuchini F.;
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of Bacillus anthracis vaccine strain 'Carbosap.'";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ413936; CAC93934.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0003405; P:pathogenesis; IEA.
DR InterPro; IPR003896; P:anthrax toxinB.
DR Pfam; PF03495; Binary_toxB; 1.
DR PRINTS; PRO1391; BINARYTOXINB.
KW Plasmid.
FT NON_TER 1 743
FT NON_TER 743 743
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5DE5A6 CRC64;
Query Match 97.7%; Score 3689; DB 2; Length 743;
Best Local Similarity 99.4%; Pred. No. 7.8e-167;
Matches 718; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 EVKQENRLNSESSESSQGLGYFSDLNQFAPMVVTSSTTGDLSIPSELENIPSENYQF 60
Db 22 EVKQENRLNSESSESSQGLGYFSDLNQFAPMVVTSSTTGDLSIPSELENIPSENYQF 81
Qy 61 QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 82 QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 141
Qy 121 QRENPTTEKGLDFKLYWTDSONKKEVISSDNQLQPELKQSSNSRKKSTASGPTVPDRN 180
Db 142 QRENPTTEKGLDFKLYWTDSONKKEVISSDNQLQPELKQSSNSRKKSTASGPTVPDRN 201
Qy 181 DGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSSPEKWSSTASDPSDFEKT 240
Db 202 DGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSSPEKWSSTASDPSDFEKT 261
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Db 262 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 321
Qy 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 322 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 381
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420
Db 382 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILLAPNNYPSKNLAPIA 441
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
Db 442 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 501
Qy 481 LPQIQTETARIIIFNGKDLNVERRIAANVPSDPLETTPDWTLEKALKIAFGFNEPKNL 540
Db 502 LPQIQTETARIIIFNGKDLNVERRIAANVPSDPLETTPDWTLEKALKIAFGFNEPKNL 561
Qy 541 QYQKDIETEFDFNFDOQTSONIKQNLAEALNATYITVLDKIKLNAMNILLIRDKRFHYDR 600
Db 562 QYQKDIETEFDFNFDOQTSONIKQNLAEALNATYITVLDKIKLNAMNILLIRDKRFHYDR 621
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLINIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 622 NNIAGADESVVKEAHEVINSSTEGLLINIDKIRKILSGYIVEIEDTEGLKEVINDRY 681
Qy 661 DMLNLSLRDQGTFFIDFKYNDKPLYSNPNYKVVVAVTKENTIIINPSENGDTSTNG 720
Db 682 DMLNLSLRDQGTFFIDFKYNDKPLYSNPNYKVVVAVTKENTIIINPSENGDTSTNG 741
Qy 721 IK 722
Db 742 IK 743
RESULT 3
Q93739 ID Q93739 PRELIMINARY; PRT; 876 AA.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 34.6261 Seconds
(without alignments)
6697.424 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENLLNSESQGLL.....TSTNGIKILFSSKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organalle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3689	97.7	741	2 Q937W2	Q937W2 bacillus an
2	3689	97.7	743	2 Q937W3	Q937W3 bacillus an
3	882.5	23.4	876	2 Q32739	Q32739 clostridium
4	882.5	23.4	879	2 Q06498	Q06498 clostridium
5	880.5	23.3	876	2 Q9KH41	Q9KH41 clostridium
6	876	23.2	875	2 Q46221	Q46221 clostridium
7	818	21.7	721	2 Q86171	Q86171 clostridium
8	728	19.3	775	2 Q844J8	Q844J8 bacillus th
9	296.5	7.9	225	2 Q8KYK2	Q8KYK2 bacillus an
10	220.5	5.8	1881	16 Q8RGK2	Q8RGK2 fusobacteri
11	216.5	5.7	3468	5 Q8II04	Q8II04 plasmodium
12	206	5.5	4688	16 Q9PQ08	Q9PQ08 ureaplasma
13	203.5	5.4	1904	5 Q8IKS2	Q8IKS2 plasmodium
14	200.5	5.3	2849	5 Q8IHY4	Q8IHY4 plasmodium
15	197.5	5.2	2940	5 Q8IHP9	Q8IHP9 plasmodium
16	195	5.2	1404	2 Q8RJN9	Q8RJN9 mycoplasma

17	192.5	5.1	675	16 Q897H6	Q897H6 clostridium
18	192.5	5.1	1301	5 Q8WSK5	Q8WSK5 plasmodium
19	190.5	5.0	2771	5 Q26216	Q26216 plasmodium
20	189.5	5.0	2444	16 Q81GX1	Q81GX1 bacillus ce
21	189	5.0	2792	5 Q814R2	Q814R2 plasmodium
22	189	5.0	9904	16 Q8NWQ6	Q8NWQ6 staphylococ
23	188	5.0	3248	5 Q8ICP9	Q8ICP9 plasmodium
24	187.5	5.0	1301	5 Q8IHQ2	Q8IHQ2 plasmodium
25	187.5	5.0	1681	5 Q8I548	Q8I548 plasmodium
26	186.5	4.9	1182	16 Q8EW90	Q8EW90 mycoplasma
27	186.5	4.9	2235	5 Q8IB09	Q8IB09 plasmodium
28	186	4.9	4261	5 Q8IFP4	Q8IFP4 plasmodium
29	185	4.9	3504	5 Q8II45	Q8II45 plasmodium
30	184.5	4.9	1125	16 Q88PM9	Q88PM9 mycoplasma
31	184.5	4.9	2529	16 Q25579	Q25579 helicobacte
32	184	4.9	1859	5 Q8IC27	Q8IC27 plasmodium
33	184	4.9	3203	5 Q7YWE7	Q7YWE7 plasmodium
34	184	4.9	3203	5 Q7YWE6	Q7YWE6 plasmodium
35	183	4.8	1326	5 Q8I5A6	Q8I5A6 plasmodium
36	183	4.8	2747	5 Q9BJX9	Q9BJX9 plasmodium
37	182.5	4.8	1387	5 Q9GZ76	Q9GZ76 plasmodium
38	182.5	4.8	3080	5 Q7YWF0	Q7YWF0 plasmodium
39	182	4.8	6713	16 Q99U54	Q99U54 staphylococ
40	182	4.8	6713	16 Q931R6	Q931R6 staphylococ
41	181	4.8	2369	5 Q26223	Q26223 plasmodium
42	181	4.8	4152	2 Q9ZHL3	Q9ZHL3 haemophilus
43	181	4.8	4152	16 Q7VLE8	Q7VLE8 haemophilus
44	181	4.8	10061	5 Q8I3Z1	Q8I3Z1 plasmodium
45	180.5	4.8	149	2 Q9RM77	Q9RM77 clostridium

ALIGNMENTS

RESULT 1

Q937W2 PRELIMINARY; PRT; 741 AA.
 ID Q937W2
 AC Q937W2
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Pag protein (Fragment).
 GN PAG.
 OS Bacillus anthracis.
 OG Plasmid pX01.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ferrara;
 RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
 RA Fasanella A., Francia M., Ciuchini F.,
 RT "Sequence analysis of the genes encoding for the major virulence
 RT factors of bacillus anthracis vaccine strain 'Carbosap'.",
 RL Submitted (Sep-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ413937; CAC93935.1;
 DR GO; GO:000576; C:extracellular; IEA.
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 DR GO; GO:0015070; F:toxin activity; IEA.
 DR GO; GO:0009405; P:pathogenesis; IEA.
 DR InterPro; IPR003896; Anthrax toxinB.
 DR Pfam; PF03495; B:Binary COXB; I.
 DR PRINTS; PR01391; BINARYTOXINE.
 KW Plasmid.
 FT NON TER 1 741
 FT NON TER 741 741
 SQ SEQUENCE 741 AA; 83153 MW; C7F95820B73065C0 CRC64;

Query Match 97.7%; Score 3689; DB 2; Length 741;

Best Local Similarity 99.4%; Pred. No. 7.8e-167;

Matches 718; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENLLNSESQGLLYGVFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:26:27 ; Search time 49.1208 Seconds
(without alignments)
4227.791 Million cell updates/sec

Title: US-09-848-909A-8

Perfect score: 3773

Sequence: 1 EVKQENRLNSESSESSQGL.....TSTNGIKILIFSKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 26Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3773	100.0	735	5 AAM51491	Aam51491 Anthrax P
2	3769	99.9	735	5 AAM51490	Aam51490 Anthrax P
3	3768	99.9	735	5 AAM51489	Aam51489 Anthrax P
4	3767	99.8	735	2 AAR60179	Aar60179 Protectiv
5	3767	99.8	735	5 AAM51483	Aam51483 Anthrax P
6	3767	99.8	735	5 AAM51488	Aam51488 Anthrax P
7	3767	99.8	735	5 AAM51493	Aam51493 Anthrax P
8	3767	99.8	736	3 AAY56959	Aay56959 B. anthra
9	3767	99.8	763	3 AAY56958	Aay56958 B. anthra
10	3767	99.8	764	3 AAB47306	Aab47306 Wild type
11	3767	99.8	764	7 ADE65872	Ade65872 Bacillus
12	3764	99.8	735	5 AAE18289	Aae18289 Bacillus
13	3764	99.8	764	6 AAE35717	Aae35717 Bacillus
14	3764	99.8	735	5 AAM51487	Aam51487 Anthrax P
15	3763	99.7	735	5 AAM51484	Aam51484 Anthrax P
16	3761	99.7	735	5 AAM51485	Aam51485 Anthrax P
17	3761	99.7	735	5 AAM51486	Aam51486 Anthrax P
18	3759	99.6	735	5 AAM51495	Aam51495 Anthrax P
19	3759	99.6	735	5 AAM51492	Aam51492 Anthrax P
20	3759	99.6	735	5 AAM51500	Aam51500 Anthrax P
21	3758	99.6	735	5 AAM51499	Aam51499 Anthrax P
22	3758	99.6	735	5 AAM51494	Aam51494 Anthrax P
23	3754	99.5	735	6 ABP71693	Abp71693 B. anthra
24	3752	99.4	764	6 AAR60183	Aar60183 PA(1-725)
25	3722.5	98.7	903	2 AAR60183	Aar60183 PA(1-725)

ALIGNMENTS

RESULT 1

AAM51491
ID AAM51491 standard; protein; 735 AA.

XX AC AAM51491;

XX AC AAM51491;

DT 01-FEB-2002 (first entry)

DE Anthrax PA mutant D425K.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX Bacillus anthracis.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 425

FT /note= "Wild-type Asp substituted by Lys"

PN WO200182788-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

XX (HARD) HARVARD COLLEGE.

XX PI Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

XX Protecting humans against anthrax using mutant B groups (anthrax

XX protective antigens) of the pore-forming binary A-B toxin of Bacillus

XX anthracis.

XX Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of
pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
moiety is anthrax protective antigen (PA) and using these mutants or
compositions of them for protecting against Bacillus anthracis infections
in humans, especially as vaccines. Note: The present sequence is not
given in the specification but is derived from the Bacillus anthracis
wild-type PA protein sequence shown in figure 13 (AAM51483)

26	3621	96.0	711	5 AAM52113	Aam52113 Anthrax P
27	3613	95.8	711	5 AAM51498	Aam51498 Anthrax P
28	3613	95.8	711	5 AAM51496	Aam51496 Anthrax P
29	3607	95.6	711	5 AAM51497	Aam51497 Anthrax P
30	3590	95.1	719	2 AAR60193	Aar60193 Modified
31	3051	80.9	595	3 AAE18288	Aae18288 Bacillus
32	2912	77.2	569	3 AAY56961	Aay56961 B. anthra
33	2502	66.3	487	5 AAE18285	Aae18285 Bacillus
34	2187	58.0	426	5 AAE18287	Aae18287 Bacillus
35	1638	43.4	318	5 AAE18286	Aae18286 Bacillus
36	1338	35.5	258	5 AAE18284	Aae18284 Bacillus
37	1275	33.8	288	5 AAM50707	Aam50707 Bacillus
38	848.5	22.5	1052	4 AAE07903	Aae07903 C. botuli
39	847	22.4	1032	4 AAE07901	Aae07901 C. botuli
40	843	22.3	1092	4 AAE07900	Aae07900 C. botuli
41	834.5	22.1	1112	4 AAE07902	Aae07902 C. botuli
42	819	21.7	721	6 AAE35719	Aae35719 Clostridi
43	781.5	20.7	880	2 AAW60224	Aaw60224 Bacillus
44	781	20.7	881	3 AAY59277	Aay59277 M1S toxin
45	780	20.7	884	2 AAR91239	Aar91239 B. cereus

481 LPOIETTARLIENGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIARFGNEPNGNL 540
481 LPOIETTARLIENGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIARFGNEPNGNL 540
541 QYQKDIETEFNFQOQTSQNIKNQLAELNATNIYTVLDKIKNAKWNILIRDKRFHYDR 600
541 QYQKDIETEFNFQOQTSQNIKNQLAELNATNIYTVLDKIKNAKWNILIRDKRFHYDR 600
601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKQIRKILSGVIVEIDTEGLKEVINDRY 660
601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKQIRKILSGVIVEIDTEGLKEVINDRY 660
661 DMLNISSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDSTNG 720
661 DMLNISSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDSTNG 720
721 IKKILIFSKKGYEIG 735
721 IKKILIFSKKGYEIG 735

RESULT 3
ID AAM51489 standard; protein; 735 AA.
XX AAM51489;
XX AC
XX 01-FEB-2002 (first entry)
XX Anthrax PA mutant D425N.
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; muten.
XX Bacillus anthracis.
XX Synthetic.
XX Key Location/Qualifiers
XX Misc-difference /note= "Wild-type Asp substituted by Asn"
XX WO200182788-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US014372.
XX 04-MAY-2000; 2000US-0201800P.
XX (HARD) HARVARD COLLEGE.
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX Claim 4; Page; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAM51483)
XX Sequence 735 AA;
SQ

Query Match 99.9%; Score 3768; DB 5; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.5e-244;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLLYGYFSDLPQAPMVTSSITGDLSPSSSELENI PSENOYF 60
DB 1 EVKQENRLNESSESSQGLLYGYFSDLPQAPMVTSSITGDLSPSSSELENI PSENOYF 60
QY 61 QSAIWSGFIKVKKSDBYTATSADNHVTWVDDQEVINKASNSNKIRLEKGLYQIKQY 120
DB 61 QSAIWSGFIKVKKSDBYTATSADNHVTWVDDQEVINKASNSNKIRLEKGLYQIKQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPKOKSSNSRKRKSTSACTVPDRON 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPKOKSSNSRKRKSTSACTVPDRON 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPSDEKVT 240
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPSDEKVT 240
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYVNOFLEKTKQLRLDTQVYGNATYVNFENGVRVDTGSNWSEV 480
DB 421 LNAQKDFSSPTITWYVNOFLEKTKQLRLDTQVYGNATYVNFENGVRVDTGSNWSEV 480
QY 481 LPQIETTARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIARFGNEPNGNL 540
DB 481 LPQIETTARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIARFGNEPNGNL 540
QY 541 QYQKDIETEFNFQOQTSQNIKNQLAELNATNIYTVLDKIKNAKWNILIRDKRFHYDR 600
DB 541 QYQKDIETEFNFQOQTSQNIKNQLAELNATNIYTVLDKIKNAKWNILIRDKRFHYDR 600
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKQIRKILSGVIVEIDTEGLKEVINDRY 660
DB 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKQIRKILSGVIVEIDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 4
ID AAR60179 standard; protein; 735 AA.
XX AAR60179;
XX AC
XX 25-MAR-2003 (revised)
XX 03-APR-1995 (first entry)
XX Protective antigen of Bacillus anthracis.
XX Anthrax; Bacillus anthracis; fusion protein; protective antigen;
XX protective antigen; cell killing; targeting; targeting; pathogen;
XX intracellular; HIV; human immunodeficiency virus; toxin.
XX Bacillus anthracis.
OS

```
XX W09418332-A2.
XX PD 18-AUG-1994.
XX PF 14-FEB-1994; 94WO-US001624.
XX PR 12-FEB-1993; 93US-00021601.
XX PR 25-JUN-1993; 93US-00082849.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Leppia SH, Klimpel K, Arora N, Singh Y, Nichols PJ;
XX WPI; 1994-279753/34.
XX DR N-PSDB; AAQ70180.
XX
PT Nucleic acid encoding anthrax toxin fusion protein - useful for
PT targeting toxin to specific cells, eg for killing tumour cells or HIV-
PT infected cells.
XX
XX PS Disclosure; Page 81-83; 124pp; English.
XX
XX CC The sequence encoding the protective antigen of Bacillus anthracis may be
XX used in the construction of a nucleic acid which encodes a fusion protein
XX comprising the anthrax protective antigen binding domain of the native
XX anthrax lethal factor and a sequence encoding an activity inducing domain
XX of a second protein. The fusion proteins are useful for the specific
XX killing of tumour cells or the killing of cells infected with
XX intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to
XX correct FN field.)
XX
XX SQ Sequence 735 AA;
XX
XX Query Match 99.8%; Score 3767; DB 2; Length 735;
XX Best Local Similarity 99.9%; Pred. No. 1.7e-244;
XX Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDVETFTATSADNHVTMVDQEVINKASNKIRLEKRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDVETFTATSADNHVTMVDQEVINKASNKIRLEKRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKKRSTSGPTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKKRSTSGPTVPDRDN 180
QY 181 DGIPTDSEVGYTVDVKNKTELSPMISNTHKXGLTKYKSSPEKSTASDPYSDPEKYT 240
DB 181 DGIPTDSEVGYTVDVKNKTELSPMISNTHKXGLTKYKSSPEKSTASDPYSDPEKYT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNEVHASFDDIGSVSAGFSNSSTVAIDHSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNEVHASFDDIGSVSAGFSNSSTVAIDHSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPYNYVLPPTSLVLGKNOIATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYNTGTAPYNYVLPPTSLVLGKNOIATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQDSSSTPTWNNYQFLEKTKOLRLDQDVYGNATYFNGRVRVDTGGSNNSEV 480
DB 421 LNAQDSSSTPTWNNYQFLEKTKOLRLDQDVYGNATYFNGRVRVDTGGSNNSEV 480
QY 481 LPOICETARTIIFNGKOLNVERRIAANVPSDPLETTKPDWTLKEALKIAFGNENGNL 540
DB 481 LPOICETARTIIFNGKOLNVERRIAANVPSDPLETTKPDWTLKEALKIAFGNENGNL 540
```

```
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAXKNILIRKRFHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAXKNILIRKRFHYDR 600
QY 601 NNIAVGADESIVKEAHREVINSSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESIVKEAHREVINSSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPYKVNVAVTYKNTIINPSENGDSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPYKVNVAVTYKNTIINPSENGDSTNG 720
QY 721 IKKILIFSCKGYEIG 735
DB 721 IKKILIFSCKGYEIG 735
XX
XX RESULT 5
XX ID AAM51483 standard; protein; 735 AA.
XX AC AAM51483;
XX DT 01-FEB-2002 (first entry)
XX DE Anthrax PA protein.
XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.
XX OS Bacillus anthracis.
XX PN WO200182788-A2.
XX PD 08-NOV-2001.
XX PF 04-MAY-2001; 2001WO-US014372.
XX PR 04-MAY-2000; 2000US-0201800P.
XX PA (HARD ) HARVARD COLLEGE.
XX PI Collier RJ, Sellman BR;
XX DR WPI; 2002-017725/02.
XX N-PSDB; AA199904.
XX PT Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX PS Disclosure; Fig 13; 77pp; English.
XX
XX CC The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. The present sequence is that of the
XX anthrax PA protein
XX
XX SQ Sequence 735 AA;
XX
XX Query Match 99.8%; Score 3767; DB 5; Length 735;
XX Best Local Similarity 99.9%; Pred. No. 1.7e-244;
XX Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDVETFTATSADNHVTMVDQEVINKASNKIRLEKRLYQIKIY 120
```

Db 61 QSAIWSGFIKVKSDSEYFATADNHNVTWVDQVINKASNKIRLEKGLVQIKIY 120
QY 121 QRENTEKGLDFKLWYDSQNKKEVSSDNLOLPELKOKSNSRKRKSTAGTPVDRDN 180
Db 121 QRENTEKGLDFKLWYDSQNKKEVSSDNLOLPELKOKSNSRKRKSTAGTPVDRDN 180
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNNVLTSLVGLKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVLTSLVGLKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNIAFYNGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNIAFYNGRVRVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNAMNILLIRDRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNAMNILLIRDRFHYDR 600
QY 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEDETEGLKEVINDRY 660
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 6
AAM51488
ID AAM51488 standard; protein; 735 AA.
XX
AC AAM51488;
XX
DT 01-FEB-2002 (first entry)
XX
DE Anthrax PA mutant D425A.
XX
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.
XX
OS Bacillus anthracis.
OS Synthetic.
PH Key Location/Qualifiers
FT Misc-difference 425 /note= "Wild-type Asp substituted by Ala"
XX
XX WO200182788-A2.
XX
PD 08-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.
XX (HARD) HARVARD COLLEGE.
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX Claim 4; Page; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: the present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAM51483)
XX Sequence 735 AA;
XX
XX Query Match 99.8%; Score 3767; DB 5; Length 735;
XX Best Local Similarity 99.9%; Pred. No. 1.7e-244;
XX Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENLLNNESSSSOGLLGYYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENLLNNESSSSOGLLGYYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDSEYFATADNHNVTWVDQVINKASNKIRLEKGLVQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYFATADNHNVTWVDQVINKASNKIRLEKGLVQIKIY 120
QY 121 QRENTEKGLDFKLWYDSQNKKEVSSDNLOLPELKOKSNSRKRKSTAGTPVDRDN 180
Db 121 QRENTEKGLDFKLWYDSQNKKEVSSDNLOLPELKOKSNSRKRKSTAGTPVDRDN 180
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNNVLTSLVGLKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVLTSLVGLKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNIAFYNGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNIAFYNGRVRVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNAMNILLIRDRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNAMNILLIRDRFHYDR 600
QY 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEDETEGLKEVINDRY 660

661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVVYAVTKENTIINPSENGDSTNG 720
 QY
 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVVYAVTKENTIINPSENGDSTNG 720
 Db
 721 IKKILIFSKGYEIG 735
 QY
 721 IKKILIFSKGYEIG 735
 Db
 RESULT 7
 AAMS1493 standard; protein; 735 AA.
 ID AAMS1493
 XX
 AC AAMS1493;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Anthrax PA mutant K397D/D425K.
 XX
 KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
 XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
 XX
 OS Bacillus anthracis.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"
 FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"
 FT
 XX W0200182788-A2.
 XX
 PD 08-NOV-2001.
 XX
 XX 04-MAY-2001; 2001WO-US014372.
 XX
 XX 04-MAY-2000; 2000US-0201800P.
 XX
 XX (HARD) HARVARD COLLEGE.
 XX
 XX Collier RJ, Sellman BR;
 XX
 XX WPI; 2002-017725/02.
 XX
 XX Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.
 XX
 PS Claim 4; Page; 77pp; English.
 XX
 XX The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections
 CC in humans, especially as vaccines. Note: the present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAMS1483)
 XX
 SQ Sequence 735 AA;
 Query March 99.8%; Score 3767; DB 5; Length 735;
 Best Local Similarity 99.9%; Pred. No. 1.7e-244;
 Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
 Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
 QY 61 QSATWSGFIKVKSGDEYTFATSNADNHTVMWDDQEVINKASNNKIRLEKGRLYQIKY 120

Db 61 QSATWSGFIKVKSGDEYTFATSNADNHTVMWDDQEVINKASNNKIRLEKGRLYQIKY 120
 QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPFLKQKSSNSRKRKSTASGPTVDRDN 180
 Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPFLKQKSSNSRKRKSTASGPTVDRDN 180
 QY 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNHEKKGITKYKSPKXWSTASDPYSDPEKVT 240
 Db 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNHEKKGITKYKSPKXWSTASDPYSDPEKVT 240
 QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTTSKNTSTRTHT 300
 Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTTSKNTSTRTHT 300
 QY 301 SEVHGNAEVEASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
 Db 301 SEVHGNAEVEASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
 QY 361 NANIRYVNTGTAPIYVLPPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPPSKNLAPIA 420
 Db 361 NANIRYVNTGTAPIYVLPPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPPSKNLAPIA 420
 QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNATATYFNGRVEVDGTSNWSEV 480
 Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNATATYFNGRVEVDGTSNWSEV 480
 QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDELTTPKDMTLEKALKTAFGNPNENGL 540
 Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDELTTPKDMTLEKALKTAFGNPNENGL 540
 QY 541 QYQKQDITFDFNFDOQTSONIKNQLAELNATINIVLDKIKLNKXNTLIRDKRPHYDR 600
 Db 541 QYQKQDITFDFNFDOQTSONIKNQLAELNATINIVLDKIKLNKXNTLIRDKRPHYDR 600
 QY 601 NNIAVGADSVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
 Db 601 NNIAVGADSVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
 QY 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVVYAVTKENTIINPSENGDSTNG 720
 Db 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVVYAVTKENTIINPSENGDSTNG 720
 QY 721 IKKILIFSKGYEIG 735
 Db 721 IKKILIFSKGYEIG 735
 RESULT 8
 AAY56959
 ID AAY56959 standard; protein; 736 AA.
 XX
 AC AAY56959;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE B. anthracis MAT-PA protein.
 XX
 KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;
 KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.
 OS
 OS Bacillus anthracis.
 XX
 PN W0200002522-A2.
 XX
 PD 20-JAN-2000.
 XX
 PF 09-JUL-1999; 99WO-US015568.
 XX
 PR 10-JUL-1998; 98US-0092416P.
 XX
 PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
 XX
 PI Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;

XX WPI; 2000-182165/16.
DR N-PSDB; AAZ56875.
XX Recombinant DNA construct useful as vaccines for anthrax, in producing
PT host cells for analyzing the drugs and agents inhibiting anthrax.
XX
PS Disclosure; Page 34; 35pp; English.
XX
CC The invention provides a recombinant DNA construct that comprises a
CC vector and at least one nucleic acid (or its fragment) encoding a
CC combination of Bacillus anthracis proteins, selected from protective
CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA
CC with its secretory signals replaced with those of tissue plasminogen
CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine
CC for anthrax and in producing infectious alpha virus particles. These
CC particles, expressing the B. anthracis proteins are useful also as
CC vaccines for anthrax. Host cells transformed with the construct are
CC useful for analyzing the effectiveness of drugs and agents that inhibit
CC anthrax or B. anthracis proteins. The present sequence represents a B.
CC anthracis MAT-PA protein
XX
SQ Sequence 736 AA;

Query Match 99.8%; Score 3767; DB 3; Length 736;
Best Local Similarity 99.9%; Pred. No. 1.7e-244;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
DB 2 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 61

QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 62 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 121

QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRSTSGPTVPDRDN 180
DB 122 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRSTSGPTVPDRDN 181

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPPEKWSASDPSDFEKT 240
DB 182 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPPEKWSASDPSDFEKT 241

QY 241 GRIDKNVSEARPLVAAPVIVHVDENILSKNEDQSTQNTDSETRTSKNTSRTHT 300
DB 242 GRIDKNVSEARPLVAAPVIVHVDENILSKNEDQSTQNTDSETRTSKNTSRTHT 301

QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 360
DB 302 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 361

QY 361 NANIRYVNTGTAPIYVLPFTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 362 NANIRYVNTGTAPIYVLPFTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 421

QY 421 LNAQKQSPPTIMYNNQFLEKTKQLRLDTPQVGNATYNFENGRVVRVDTGNSWSEV 480
DB 422 LNAQKQSPPTIMYNNQFLEKTKQLRLDTPQVGNATYNFENGRVVRVDTGNSWSEV 481

QY 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPENGL 540
DB 482 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPENGL 541

QY 541 QYQKQITTEFDNFDDQTSQNIKNQALAEANATNIVYVLDKIKLNKQNNILIRDKRPHYDR 600
DB 542 QYQKQITTEFDNFDDQTSQNIKNQALAEANATNIVYVLDKIKLNKQNNILIRDKRPHYDR 601

QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660
DB 602 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY 661

QY 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720

DB 662 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 721
QY 721 IKKILIFSCKGYEIG 735
DB 722 IKKILIFSCKGYEIG 736

RESULT 9
AAZ56960
ID AAZ56960 standard; protein; 763 AA.
XX
XX AAZ56960;
XX
DT 25-APR-2000 (first entry)
XX
DE B. anthracis TPA-PA protein.
XX
KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;
KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.
XX
OS Bacillus anthracis.
XX
XX WO200002522-A2.
XX
PD 20-JAN-2000.
XX
PF 09-JUL-1999; 99WO-US015568.
XX
PR 10-JUL-1998; 98US-0092416P.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
XX Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;
PI WPI; 2000-182165/16.
DR N-PSDB; AAZ56876.
XX
PT Recombinant DNA construct useful as vaccines for anthrax, in producing
PT host cells for analyzing the drugs and agents inhibiting anthrax.
XX
PS Disclosure; Page 32; 35pp; English.
XX
CC The invention provides a recombinant DNA construct that comprises a
CC vector and at least one nucleic acid (or its fragment) encoding a
CC combination of Bacillus anthracis proteins, selected from protective
CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA
CC with its secretory signals replaced with those of tissue plasminogen
CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine
CC for anthrax and in producing infectious alpha virus particles. These
CC particles, expressing the B. anthracis proteins are useful also as
CC vaccines for anthrax. Host cells transformed with the construct are
CC useful for analyzing the effectiveness of drugs and agents that inhibit
CC anthrax or B. anthracis proteins. The present sequence represents a B.
CC anthracis TPA-PA protein
XX
SQ Sequence 763 AA;

Query Match 99.8%; Score 3767; DB 3; Length 763;
Best Local Similarity 99.9%; Pred. No. 1.8e-244;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
DB 29 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 88

QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 89 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 148

QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRSTSGPTVPDRDN 180
DB 149 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRSTSGPTVPDRDN 208

QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYSDFEKVT 240
 Db 209 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYSDFEKVT 268
 QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQMTDSETRISKVTSRTH 300
 Db 269 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQMTDSETRISKVTSRTH 328
 QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
 Db 329 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 388
 QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
 Db 389 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 448
 QY 421 LNAQKDFSTPTIMWYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSWSEV 480
 Db 449 LNAQKDFSTPTIMWYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSWSEV 508
 QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
 Db 509 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 568
 QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMILLRDKRFHYDR 600
 Db 569 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMILLRDKRFHYDR 628
 QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
 Db 629 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 688
 QY 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNKYVNYAVTKENTIINPSENGDTSTNG 720
 Db 689 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNKYVNYAVTKENTIINPSENGDTSTNG 748
 QY 721 IKKILIFSKKGYEIG 735
 Db 749 IKKILIFSKKGYEIG 763
 RESULT 10
 AAY56958
 ID AAY56958 standard; protein: 764 AA.
 XX
 AC AAY56958;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE B. anthracis protective antigen (PA) protein.
 XX
 KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;
 KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.
 XX
 OS Bacillus anthracis.
 XX
 FN WO200002522-A2.
 XX
 PD 20-JAN-2000.
 XX
 PF 09-JUL-1999; 99WO-US015568.
 XX
 PR 10-JUL-1998; 98US-0092416P.
 XX
 PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
 XX
 PI Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;
 XX
 DR WPI; 2000-182165/16.
 DR N-PSDB; AAZ56874.
 XX
 PT Recombinant DNA construct useful as vaccines for anthrax, in producing

PT host cells for analyzing the drugs and agents inhibiting anthrax.
 XX Disclosure; Page 33; 35pp; English.
 XX
 CC The invention provides a recombinant DNA construct that comprises a
 CC vector and at least one nucleic acid (or its fragment) encoding a
 CC combination of Bacillus anthracis proteins, selected from protective
 CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA
 CC with its secretory signals replaced with those of tissue plasminogen
 CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine
 CC for anthrax and in producing infectious alpha virus particles. These
 CC particles, expressing the B. anthracis proteins are useful also as
 CC vaccines for anthrax. Host cells transformed with the construct are
 CC useful for analyzing the effectiveness of drugs and agents that inhibit
 CC anthrax or B. anthracis proteins. The present sequence represents a B.
 CC anthracis PA protein
 XX
 SQ Sequence 764 AA;
 Query Match 99.8%; Score 3767; DB 3; Length 764;
 Best Local Similarity 99.9%; Pred. No. 1.8e-244;
 Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EVQENRLLNESSESSQGLLGYYFSDLNPOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
 Db 30 EVQENRLLNESSESSQGLLGYYFSDLNPOAPMVVTSSTGDLSPSSSELENIPSENQYF 89
 QY 61 QSAIWSGFIKVKSDSEYTPATSADNHVTVMVDDQEVINKASNSKIRLEKRLYQIKQY 120
 Db 90 QSAIWSGFIKVKSDSEYTPATSADNHVTVMVDDQEVINKASNSKIRLEKRLYQIKQY 149
 QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNSRKRKRSAGTSVDPDRN 180
 Db 150 QRENPTKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNSRKRKRSAGTSVDPDRN 209
 QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYSDFEKVT 240
 Db 210 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYSDFEKVT 269
 QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQMTDSETRISKVTSRTH 300
 Db 270 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQMTDSETRISKVTSRTH 329
 QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
 Db 330 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 389
 QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
 Db 390 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 449
 QY 421 LNAQKDFSTPTIMWYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSWSEV 480
 Db 450 LNAQKDFSTPTIMWYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSWSEV 509
 QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
 Db 510 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 569
 QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMILLRDKRFHYDR 600
 Db 570 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMILLRDKRFHYDR 629
 QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
 Db 630 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 689
 QY 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNKYVNYAVTKENTIINPSENGDTSTNG 720
 Db 690 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNKYVNYAVTKENTIINPSENGDTSTNG 749
 QY 721 IKKILIFSKKGYEIG 735
 |||||

Db	750 IKKILIFSKGYEIG 764	30 EVKQENRLNSESSESSQGLLYYFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQVNF 89
Db	750 IKKILIFSKGYEIG 764	61 QSAIWSGFIKVKSDYVTFATSDADNHVTWVDDQBYINKASNSNKLRLKGRLYQIKIY 120
Db	750 IKKILIFSKGYEIG 764	90 QSAIWSGFIKVKSDYVTFATSDADNHVTWVDDQBYINKASNSNKLRLKGRLYQIKIY 149
Db	750 IKKILIFSKGYEIG 764	121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNQLPELQKXSSNSRKRKSTAGTVPDRDN 180
Db	750 IKKILIFSKGYEIG 764	150 QRENPTKGLDFKLYWTDSONKKEVIVSSDNQLPELQKXSSNSRKRKSTAGTVPDRDN 209
Db	750 IKKILIFSKGYEIG 764	181 DGPDSLEVEGYTVDVKNKTFELSPWISNHEKKGTLTKYKSSPEKSTASDPYDEKVT 240
Db	750 IKKILIFSKGYEIG 764	210 DGPDSLEVEGYTVDVKNKTFELSPWISNHEKKGTLTKYKSSPEKSTASDPYDEKVT 269
Db	750 IKKILIFSKGYEIG 764	241 GRIDKNVSPARHPLVAAVPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTRTHT 300
Db	750 IKKILIFSKGYEIG 764	270 GRIDKNVSPARHPLVAAVPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTRTHT 329
Db	750 IKKILIFSKGYEIG 764	301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Db	750 IKKILIFSKGYEIG 764	330 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 389
Db	750 IKKILIFSKGYEIG 764	361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db	750 IKKILIFSKGYEIG 764	390 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 449
Db	750 IKKILIFSKGYEIG 764	421 LNAQKDFSTPTITMNYNQFLEKTKQLRLDTPQVYGNIAIYFENGSRVVDGTSNWEV 480
Db	750 IKKILIFSKGYEIG 764	450 LNAQKDFSTPTITMNYNQFLEKTKQLRLDTPQVYGNIAIYFENGSRVVDGTSNWEV 509
Db	750 IKKILIFSKGYEIG 764	481 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPNGNL 540
Db	750 IKKILIFSKGYEIG 764	510 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPNGNL 569
Db	750 IKKILIFSKGYEIG 764	541 QYQKDIETFDNPDQOTSONIKNQLAELNATNIYTVLDKIKLNAKNNILIRKRFHYDR 600
Db	750 IKKILIFSKGYEIG 764	570 QYQKDIETFDNPDQOTSONIKNQLAELNATNIYTVLDKIKLNAKNNILIRKRFHYDR 629
Db	750 IKKILIFSKGYEIG 764	601 NNIAVGADESUVKHAHREVINSSTEGILLNIDKIRKILSGYIIVEIEDTEGLKEVINDRY 660
Db	750 IKKILIFSKGYEIG 764	630 NNIAVGADESUVKHAHREVINSSTEGILLNIDKIRKILSGYIIVEIEDTEGLKEVINDRY 689
Db	750 IKKILIFSKGYEIG 764	661 DMLMISSLRQDKTFIDFKYNDKLPYISNPNTKVNVAVTAKNTIINFSENGDTSTNG 720
Db	750 IKKILIFSKGYEIG 764	690 DMLMISSLRQDKTFIDFKYNDKLPYISNPNTKVNVAVTAKNTIINFSENGDTSTNG 749
Db	750 IKKILIFSKGYEIG 764	721 IKKILIFSKGYEIG 735
Db	750 IKKILIFSKGYEIG 764	750 IKKILIFSKGYEIG 764
Db	750 IKKILIFSKGYEIG 764	RESULT 12
Db	750 IKKILIFSKGYEIG 764	ADE65872
Db	750 IKKILIFSKGYEIG 764	ID ADE65872 standard; protein; 857 AA.
Db	750 IKKILIFSKGYEIG 764	XX ADE65872;
Db	750 IKKILIFSKGYEIG 764	DT 29-JAN-2004 (first entry)
Db	750 IKKILIFSKGYEIG 764	XX Bacillus anthracis PA antigen.
Db	750 IKKILIFSKGYEIG 764	DE Antibacterial; Vaccine; immune response; Bacillus anthracis.
Db	750 IKKILIFSKGYEIG 764	XX Bacillus anthracis.
Db	750 IKKILIFSKGYEIG 764	OS WO2003087378-A1.
Db	750 IKKILIFSKGYEIG 764	PN 23-OCT-2003.
Db	750 IKKILIFSKGYEIG 764	PD 11-APR-2003; 2003WO-GB001553.
Db	750 IKKILIFSKGYEIG 764	PF 11-APR-2002; 2002US-0371416P.
Db	750 IKKILIFSKGYEIG 764	XX
Db	750 IKKILIFSKGYEIG 764	XX

RESULT 11

AAB47306

ID AAB47306 standard; protein; 764 AA.

AC AAB47306;

XX 29-AUG-2001 (first entry)

DE Wild type B. anthracis protective antigen.

KW Lethal factor; LF; immunogen; LF4; protective antigen; PA; DNA vaccine;

XX humoral; cell-mediated; immune memory response.

OS Bacillus anthracis.

XX Key Location/Qualifiers

PH Peptide 1..29

FT /label= Signal peptide

FT /note= "Not given in the specification"

FT Protein 30..764

FT /label= PA

FT Peptide 204..764

FT /label= pCPA

XX WO200145639-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000WO-US034912.

XX 22-DEC-1999; 99US-0171459P.

XX (OHIS) UNIV OHIO STATE RES FOUND.

XX (GALL/) GALLOWAY D R.

XX (MATE/) MATECZUN A J.

XX Galloway DR, Mateczun AJ;

XX WPI; 2001-408540/43.

XX N-PSDB; AAC86016.

XX Protecting animal against lethal infection with Bacillus anthracis, by

PT administering wildtype or mutated form of Bacillus anthracis lethal

PT factor protein or its fragment or a nucleic acid encoding the mutated

PT protein.

XX Claim 5; Fig 2; 33pp; English.

XX This sequence shows the B. anthracis protective antigen (PA). An

XX immunogenic fragment of PA, pCPA, can be used to produce an immune

XX response which protects an animal against lethal infection with Bacillus

XX anthracis. DNA encoding the B. anthracis PA can be used in conjunction

XX with DNA encoding the lethal factor (LF) in a DNA vaccine. Using a DNA

XX vaccine which encodes the mutated LF protein or fragment alone or in

XX combination with a DNA encoding the PA protein or its fragment, both

XX components (humoral and cell-mediated) of the immune system are

XX stimulated, which results in longer term immune memory response. The

XX combined use of a mutated LF and PA gene or their fragments results in a

XX higher level of immune response, as judged by overall serum antibody

XX titers for LF and PA antigens, than the use of either LF or PA genes in

XX separate immunizations

XX Sequence 764 AA;

SQ

Query Match 99.8%; Score 3767; DB 4; Length 764;

Best Local Similarity 99.9%; Pred. No. 1.8e-244;

Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYYFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQVNF 60

|||||

Db 742 DMLNTSSLRQDGKTFIDFKKYNKCLPLYISPNYKVNVAVTKENTINPSENGDTSTNG 801

Qy 721 IKKILFSSKGYEIG 735

Db 802 IKKILFSSKGYEIG 816

RESULT 13

AAE18289

ID AAE18289 standard; protein; 735 AA.

XX AAE18289;

AC AAE18289;

DT 07-MAY-2002 (first entry)

XX Bacillus subtilis protective antigen (PA) domain 1+2+3+4.

DE Immunogenic reagent; immune response; protective antigen; PA; vaccine;

KW Bacillus anthracis infection; antibacterial.

XX Bacillus subtilis.

OS Bacillus subtilis.

XX Key Location/Qualifiers

FT Misc-difference 285 /note= "Encoded by GAA"

XX WO200204646-A1.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-GB003065.

XX 08-JUL-2000; 2000GB-00016702.

XX (MINA) UK SEC FOR DEFENCE.

XX Williamson ED, Miller J, Walker NJ, Baillie LMJ, Holden PT;

PI Flick-Smith HC, Bullifant HL, Titball RW, Topping AW;

XX WPI; 2002-171720/22.

DR N-PSDB; AAD29121.

XX New immunogenic reagent having a polypeptide of the full length

PT Protective Antigen of Bacillus anthracis, useful for treating B.

PT anthracis infection or in preparing a medicament for the prophylaxis or

PT treatment of the infection.

XX Example 5; Fig 3; 40pp; English.

XX The present invention relates to an immunogenic reagent, which produces

CC an immune response that is protective against Bacillus anthracis. The

CC reagent comprises one or more polypeptides which together represent up to

CC three domains of the full length Protective Antigen (PA) of Bacillus

CC anthracis or variants of these, and at least one of the domains comprises

CC domain 1 or domain 4 of PA or its variant. The invention is used as a

CC vaccine. The immunogenic reagent is useful in the preparation of a

CC medicament for the prophylaxis or treatment of B. anthracis infection.

CC The present sequence is Bacillus subtilis protective antigen domain

CC 1+2,3+4

XX SQ Sequence 735 AA;

Query Match 99.8%; Score 3764; DB 5; Length 735;

Best Local Similarity 99.7%; Pred. No. 2.8e-244;

Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60

Db 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60

Qy 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKY 120

Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKY 120

PA (POWD-) POWDERJECT RES LTD.

XX Schmaljohn C, Fuller J;

XX WPI; 2003-877105/81.

DR N-PSDB; AD565871.

XX New polynucleotide vaccine composition comprising a nucleic acid sequence

PT that encodes a Bacillus anthracis antigen, useful for eliciting a

PT protective immune response against Bacillus anthracis.

XX Example 1; SEQ ID NO 4; 65pp; English.

XX The present invention relates to a new polynucleotide vaccine composition

CC comprising a nucleic acid sequence that encodes a Bacillus anthracis

CC antigen and that is operatively linked to a promoter suitable for

CC expression of the antigen in a mammalian cell. The polynucleotide vaccine

CC composition is useful for eliciting a protective immune response against

CC Bacillus anthracis. The present sequence represents Bacillus anthracis PA

CC antigen.

XX SQ Sequence 857 AA;

Query Match 99.8%; Score 3767; DB 7; Length 857;

Best Local Similarity 99.9%; Pred. No. 2.2e-244;

Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60

Db 82 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 141

Qy 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKY 120

Db 142 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKY 201

Qy 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQIPELKQSSNRKSTASGTPVDPDRN 180

Db 202 QRENPTKGLDFKLYWTDSONKEVSSDNLQIPELKQSSNRKSTASGTPVDPDRN 261

Qy 181 DGIPTDSEVEGYTVDNKKTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYDPEKT 240

Db 262 DGIPTDSEVEGYTVDNKKTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYDPEKT 321

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTSKNTSTRTHT 300

Db 322 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTSKNTSTRTHT 381

Qy 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Db 382 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 441

Qy 361 NANIRYVNTGTAPYVNLPTSLVKGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420

Db 442 NANIRYVNTGTAPYVNLPTSLVKGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 501

Qy 421 LNAQKDFSTITWNYNQFLEKTKQLRLDQDVGNITATYFENGVRVVDTCGNWSEV 480

Db 502 LNAQKDFSTITWNYNQFLEKTKQLRLDQDVGNITATYFENGVRVVDTCGNWSEV 561

Qy 481 LPQIQTETARIIFNGKDLNLYVERIAAVNPSDPLETTKPDWTLKEALKIAGFNEPENG 540

Db 562 LPQIQTETARIIFNGKDLNLYVERIAAVNPSDPLETTKPDWTLKEALKIAGFNEPENG 621

Qy 541 QYQGHDXITFEPNFDOQTSQNIKNQALNATNYVLDKIKNAKNILIRDKRPHYDR 600

Db 622 QYQGHDXITFEPNFDOQTSQNIKNQALNATNYVLDKIKNAKNILIRDKRPHYDR 681

Qy 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKIDKILSGYIIVEIETDEGLKEVINDRY 660

Db 682 NNIAVGADESUVKEAHREVINSSTEGLLNIDKIDKILSGYIIVEIETDEGLKEVINDRY 741

Qy 661 DMLNTSSLRQDGKTFIDFKKYNKCLPLYISPNYKVNVAVTKENTINPSENGDTSTNG 720

PI Sutton JM, Shone CC;
XX WPI; 2003-167247/16.
XX Conjugate for modulating cell survival and cell growth, modulating
PT release of inflammatory mediator from cells, comprises injected bacterial
PT effector protein and a carrier that targets the protein to target cell.
XX
PS Example 12; Page 122-125; 130pp; English.
XX The invention relates to a conjugate comprising an injected bacterial
CC effector protein and a carrier that targets the effector protein to a
CC target cell. Pharmaceutical composition of the invention is useful for a
CC treatment selected from promoting or inhibiting survival of cells;
CC preventing and reversing damage to cells; killing cells; promoting or
CC inhibiting the growth of cells; apoptosis; release of an inflammatory
CC mediator from cells, division of cells and treating intracellular
CC infection and regulating nitric oxide release from cells. The invention
CC is useful in the manufacture of a medicament for treating a neuronal
CC cell, for intracellular infection, for interfering with intracellular
CC trafficking, for modulating expression of cell-surface markers and for
CC inhibiting secretion from cells. The invention is also useful for
CC treating Prion disease, Alzheimer' disease and wide range of disorders
CC including muscle spasms such as blepharospasm, torticollis and
CC hypersecretion disorders such as chronic obstructive pulmonary disease
CC (COPD), bronchitis and asthma. The present sequence is Bacillus anthracis
CC protective antigen. This sequence is used in the exemplification of the
CC invention
XX
XX Sequence 764 AA;
Query Match 99.8%; Score 3764; DB 6; Length 764;
Best Local Similarity 99.7%; Pred. No. 2.9e-244;
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENLLNESESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYQF 60
Db 30 EVKQENLLNESESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYQF 89
QY 61 QSAIWSGFIVKVKSDGYTFATSDADNHYTMVDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 90 QSAIWSGFIVKVKSDGYTFATSDADNHYTMVDDQEVINKASNNKIRLEKGRLYQIKIY 149
QY 121 QRENTEKGLDFKLYWTDSONKKEVVISDNLQPELKQKSSNRKKESTAGTVPDRDN 180
Db 150 QRENTEKGLDFKLYWTDSONKKEVVISDNLQPELKQKSSNRKKESTAGTVPDRDN 209
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSTASDPYDFEKT 240
Db 210 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSTASDPYDFEKT 269
QY 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSEITISKNITSTSRHT 300
Db 270 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSEITISKNITSTSRHT 329
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITGLNTADTARL 360
Db 330 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITGLNTADTARL 389
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 390 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 449
QY 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGNNWSEV 480
Db 450 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGNNWSEV 509
QY 481 LPQIOETTARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPQNL 540
Db 510 LPQIOETTARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPQNL 569
QY 541 QYQKDIITEFDNFDOQTSONIKOLAEINATNYTVLDKIKLNAKONILLRDKRPHYDR 600

QY 121 QRENTEKGLDFKLYWTDSONKKEVVISDNLQPELKQKSSNRKKESTAGTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVVISDNLQPELKQKSSNRKKESTAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSEITISKNITSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSEITISKNITSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGNNWSEV 480
Db 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGNNWSEV 480
QY 481 LPQIOETTARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPQNL 540
Db 481 LPQIOETTARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPQNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKOLAEINATNYTVLDKIKLNAKONILLRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKOLAEINATNYTVLDKIKLNAKONILLRDKRPHYDR 600
QY 601 NNTAVGADESVKAEHREVNSTEGLLNIDKIRKILSGYIVEDETEGLKEVINDRY 660
Db 601 NNTAVGADESVKAEHREVNSTEGLLNIDKIRKILSGYIVEDETEGLKEVINDRY 660
QY 661 DMLNISSLRODGTFFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGTFFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIPSKKGYEIG 735
Db 721 IKKILIPSKKGYEIG 735
RESULT 14
ID AAE35717
XX AAE35717 standard; protein; 764 AA.
XX
XX AAE35717;
XX
XX 17-JUN-2003 (first entry)
XX
XX Bacillus anthracis protective antigen.
XX Apoptosis; therapy; inflammatory mediator; intracellular trafficking;
XX infection; Prion disease; Alzheimer' disease; hypersecretion disorder;
XX muscle spasm; COPD; bronchitis; chronic obstructive pulmonary disease;
XX torticollis; blepharospasm; asthma; protective antigen.
XX
XX Bacillus anthracis.
XX
XX W0200296467-A2.
XX
XX 05-DEC-2002.
XX
XX 21-MAY-2002; 2002WO-GB002384.
XX
XX 24-MAY-2001; 2001GB-00012687.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX

Db 570 QYQKDIITFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKPFHYDR 629
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 630 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 689
QY 661 DMLNIISSLRQDGKTFIDFKKNDKLPYISNPNYKVNVTAVTKENTIINPSENGDTSTNG 720
Db 690 DMLNIISSLRQDGKTFIDFKKNDKLPYISNPNYKVNVTAVTKENTIINPSENGDTSTNG 749
QY 721 IKKILIFSCKGYEIG 735
Db 750 IKKILIFSCKGYEIG 764

RESULT 15
AAM51487
ID AAM51487 standard; protein; 735 AA.
XX AC AAM51487;
XX DT 01-FEB-2002 (first entry)
XX DE Anthrax PA mutant K397Q.
XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; muten.
XX OS Bacillus anthracis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 397 /note= "Wild-type Lys substituted by Gln"
XX PN W0200182788-A2.
XX PD 08-NOV-2001.
XX PF 04-MAY-2001; 2001WO-US014372.
XX PR 04-MAY-2000; 2000US-0201800P.
XX PA (HARD) HARVARD COLLEGE.
XX PI Collier RJ, Sellman BR;
XX DR WPI; 2002-017725/02.
XX PT Protecting humans against anthrax using mutant B groups (anthrax
XX PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX PT anthracis.
XX PS Claim 4; Page; 77pp; English.
XX CC The invention relates to antibacterial agents comprising mutant forms of
XX CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX CC moiety is anthrax protective antigen (PA) and using these mutants or
XX CC compositions of them for protecting against Bacillus anthracis infections
XX CC in humans, especially as vaccines. Note: The present sequence is not
XX CC given in the specification but is derived from the Bacillus anthracis
XX CC wild-type PA protein sequence shown in figure 13 (AAM51483)
XX SQ Sequence 735 AA;

Query Match 99.7%; Score 3763; DB 5; Length 735;
Best Local Similarity 99.7%; Pred. No. 3.2e-244;
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENLLNESSESSOGLIGYFSDLNFCAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENLLNESSESSOGLIGYFSDLNFCAPMVVTSSTTGDLSIPSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLSEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLSEKGRLYQIKIY 120
QY 121 QRENPEKGLDFKLYWTDSONKKEVLSNNIQLPELKOKSSNSRKKRSTSSAGTVPDRDN 180
Db 121 QRENPEKGLDFKLYWTDSONKKEVLSNNIQLPELKOKSSNSRKKRSTSSAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDEKVT 240
Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDEKVT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDHNIILSKNEDQSTQNTDSETRTTSKNTSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDHNIILSKNEDQSTQNTDSETRTTSKNTSTRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVLPPTSLVLGKNQTLATIKAKENOLSOILAPNNYYPKNIAPIA 420
Db 361 NANIRYNTGTAPIYVLPPTSLVLGKNQTLATIKAKENOLSOILAPNNYYPKNIAPIA 420
QY 421 LNAQKDPSSPTITWNYNQFLEKTKQLRLDTPQVYGNIAIYNFENGVRVDTGSNWEV 480
Db 421 LNAQKDPSSPTITWNYNQFLEKTKQLRLDTPQVYGNIAIYNFENGVRVDTGSNWEV 480
QY 481 LPQIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLKEALKIAFGNEPENGIL 540
Db 481 LPQIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLKEALKIAFGNEPENGIL 540
QY 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKPFHYDR 600
Db 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKPFHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNIISSLRQDGKTFIDFKKNDKLPYISNPNYKVNVTAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNIISSLRQDGKTFIDFKKNDKLPYISNPNYKVNVTAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

Search completed: May 3, 2004, 19:35:55
Job time : 52.1208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:31:43 ; Search time 15.0986 Seconds
(without alignments)
2513.152 Million cell updates/sec

Title: US-09-848-909A-8
Perfect score: 3773
Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKKILFSGKGYEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pdp.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pdp.*
3: /cgn2_6/ptodata/2/iaa/5A_COMB.pdp.*
4: /cgn2_6/ptodata/2/iaa/5B_COMB.pdp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pdp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3767	99.8	735	1	US-08-021-601-4
2	3767	99.8	735	1	US-08-082-849B-4
3	3767	99.8	735	5	PCT-US94-01624-4
4	3722.5	98.7	903	1	US-08-021-601-12
5	3722.5	98.7	903	1	US-08-082-849B-12
6	3722.5	98.7	903	5	PCT-US94-01624-12
7	3590	95.1	719	1	US-08-082-849B-31
8	3590	95.1	719	5	PCT-US94-01624-31
9	1275	33.8	288	4	US-09-273-839A-8
10	781	20.7	881	3	US-08-960-780-32
11	781	20.7	881	3	US-09-073-898-32
12	781	20.7	881	4	US-09-307-106-8
13	781	20.7	881	4	US-09-850-351A-32
14	779	20.6	884	1	US-08-471-033-5
15	779	20.6	884	1	US-08-471-044-5
16	779	20.6	884	2	US-08-463-483A-5
17	779	20.6	884	2	US-08-471-046A-5
18	779	20.6	884	2	US-08-470-566B-5
19	779	20.6	884	2	US-08-469-334-5
20	779	20.6	884	3	US-09-300-529-5
21	779	20.6	1346	1	US-08-471-033-23
22	779	20.6	1346	2	US-08-471-044-23
23	779	20.6	1346	2	US-08-463-483A-23
24	779	20.6	1346	2	US-08-471-046A-23
25	779	20.6	1346	2	US-08-470-566B-23
26	779	20.6	1346	2	US-08-469-334-23
27	779	20.6	1346	3	US-09-300-529-23

28 778 20.6 852 1 US-08-471-033-36 Sequence 36, Appl
29 778 20.6 852 2 US-08-471-044-36 Sequence 36, Appl
30 778 20.6 852 2 US-08-463-483A-36 Sequence 36, Appl
31 778 20.6 852 2 US-08-471-046A-36 Sequence 36, Appl
32 778 20.6 852 2 US-08-470-566B-36 Sequence 36, Appl
33 778 20.6 852 2 US-08-469-334-36 Sequence 36, Appl
34 778 20.6 852 3 US-09-300-529-36 Sequence 36, Appl
35 778 20.6 1338 1 US-08-471-033-50 Sequence 50, Appl
36 778 20.6 1338 2 US-08-471-044-50 Sequence 50, Appl
37 778 20.6 1338 2 US-08-463-483A-50 Sequence 50, Appl
38 778 20.6 1338 2 US-08-471-046A-50 Sequence 50, Appl
39 778 20.6 1338 2 US-08-470-566B-50 Sequence 50, Appl
40 778 20.6 1338 2 US-08-469-334-50 Sequence 50, Appl
41 778 20.6 1338 3 US-09-300-529-50 Sequence 50, Appl
42 768.5 20.4 784 3 US-09-371-913A-7 Sequence 7, Appl
43 768.5 20.4 784 4 US-09-967-805-7 Sequence 7, Appl
44 759.5 20.1 860 4 US-09-307-106-48 Sequence 48, Appl
45 749.5 19.9 834 1 US-08-471-033-21 Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-021-601-4
; Sequence 4, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 133 Carnegie Way, Suite 400
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,601
FILING DATE: 19930212
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-021-601-4

Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.3e-256;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EVKQENLLNESSSQGLLGYEISLNFQAPMVVTSSTTGDLSIFSSLELIPSENQYF 60
|||||

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-4

Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.3e-256;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENLLNESESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWDSQNKKEVSSDNLOLPELKOKSSNSRKRSTAGTPVDPDRDN 180
DB 121 QRENTEKGLDFKLYWDSQNKKEVSSDNLOLPELKOKSSNSRKRSTAGTPVDPDRDN 180
QY 181 DGIPODSEVGYTVVKNKRTFLSPWISNIHEKKGTLTKYKSPKXSTASDPYSDFEKT 240
DB 181 DGIPODSEVGYTVVKNKRTFLSPWISNIHEKKGTLTKYKSPKXSTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
QY 301 SEVHGNAEVHASFFDGGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFFDGGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNYVLPPTSLVLGKNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYNYVLPPTSLVLGKNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYFNENGVRVDTGNNWSEV 480
DB 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYFNENGVRVDTGNNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
DB 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 541 QYQGDITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLAKNNILLRDKRFHYDR 600
DB 541 QYQGDITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLAKNNILLRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIIEDETEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIIEDETEGLKEVINDRY 660

DB 1 EVKQENLLNESESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWDSQNKKEVSSDNLOLPELKOKSSNSRKRSTAGTPVDPDRDN 180
DB 121 QRENTEKGLDFKLYWDSQNKKEVSSDNLOLPELKOKSSNSRKRSTAGTPVDPDRDN 180
QY 181 DGIPODSEVGYTVVKNKRTFLSPWISNIHEKKGTLTKYKSPKXSTASDPYSDFEKT 240
DB 181 DGIPODSEVGYTVVKNKRTFLSPWISNIHEKKGTLTKYKSPKXSTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
QY 301 SEVHGNAEVHASFFDGGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFFDGGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNYVLPPTSLVLGKNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYNYVLPPTSLVLGKNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYFNENGVRVDTGNNWSEV 480
DB 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYFNENGVRVDTGNNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
DB 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 541 QYQGDITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLAKNNILLRDKRFHYDR 600
DB 541 QYQGDITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLAKNNILLRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIIEDETEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIIEDETEGLKEVINDRY 660
QY 661 DMLNISRQDKGTFFDKYNDKPLYSNPKNYKVVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISRQDKGTFFDKYNDKPLYSNPKNYKVVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 2
US-082-849B-4
Sequence 4, Application US/08082849B
Parent No. 5677274
GENERAL INFORMATION:
APPLICANT: Lepplia, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
RELATED METHODS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

QY 241 GRIDKXVSPARHPVAAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKXVSPARHPVAAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTMTGLNTADTARL 360
DB 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTMTGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLELEKTQKRLDQDQVYGNATYVNFENGRVVDVTGNSWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLELEKTQKRLDQDQVYGNATYVNFENGRVVDVTGNSWSEV 480
QY 481 LPQIQETIARIIFNGKDLNVERRAAANPSDPLETTKPDWTLKALKIAPFNGPENGML 540
DB 481 LPQIQETIARIIFNGKDLNVERRAAANPSDPLETTKPDWTLKALKIAPFNGPENGML 540
QY 541 QYQKDIETEFDFNDOOTSQNIKNQALAEALNATYVVDKIKLNKQNLILIRDKRFHYDR 600
DB 541 QYQKDIETEFDFNDOOTSQNIKNQALAEALNATYVVDKIKLNKQNLILIRDKRFHYDR 600
QY 601 NNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYVIEIETEGLEKEVINDRY 660
DB 601 NNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYVIEIETEGLEKEVINDRY 660
QY 661 DMLNLSLRQDQKTFIDFKKNDKPLYSNPNYKVVAVTKENTIIINPSNGDTSTNG 720
DB 661 DMLNLSLRQDQKTFIDFKKNDKPLYSNPNYKVVAVTKENTIIINPSNGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 4

US-08-021-601-12
; Sequence 12, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klumpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 133 Carnegie Way, Suite 400
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,601
; FILING DATE: 19930212
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880

QY 661 DMLNLSLRQDQKTFIDFKKNDKPLYSNPNYKVVAVTKENTIIINPSNGDTSTNG 720
DB 661 DMLNLSLRQDQKTFIDFKKNDKPLYSNPNYKVVAVTKENTIIINPSNGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735
RESULT 3
PCT-US94-01624-4
; Sequence 4, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klumpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND KOURIE and CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01624-4

Query Match 99.8%; Score 3767; DB 5; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.3e-256;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENRLNESSSQGLLYGFFSDLNFAQPMVVTSSSTGDLSTPSSSELENIPSENYQF 60
DB 1 EVKQENRLNESSSQGLLYGFFSDLNFAQPMVVTSSSTGDLSTPSSSELENIPSENYQF 60
QY 61 QSAIWSGFTKVKSDYTTATSDADNHVTWVDDQEVINKASNKRIRLEKGLYQIKQY 120
DB 61 QSAIWSGFTKVKSDYTTATSDADNHVTWVDDQEVINKASNKRIRLEKGLYQIKQY 120
QY 121 QRENPTKGLDFKLYWTDQNKKEVTSNDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDQNKKEVTSNDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVNKKETFLSPWTSNHEKKGITKYSKSPKWKSTADPYSDPEKVT 240
DB 181 DGIPDSLEVEGYTVDVNKKETFLSPWTSNHEKKGITKYSKSPKWKSTADPYSDPEKVT 240

```
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-021-601-12

Query Match      98.7%; Score 3722.5; DB 1; Length 903;
Best Local Similarity 99.0%; Pred. No. 2.3e-253;
Matches 727; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 EVKQENRLNESSSQGLLYGFSDLNFOAPMVVTSSTTGLSPSSSELENIPSENOYF 60
DB 1 EVKQENRLNESSSQGLLYGFSDLNFOAPMVVTSSTTGLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTASDPSDFEKT 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTASDPSDFEKT 180
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKMWSTASDPSDFEKT 240
DB 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKMWSTASDPSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNYLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYNYLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWNYNQFLKTKQLRLDTQVYGNATYNFENGVRVDTGNSWSEV 480
DB 421 LNAQKDFSSPTITWNYNQFLKTKQLRLDTQVYGNATYNFENGVRVDTGNSWSEV 480
QY 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNPNGNL 540
DB 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNPNGNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKKNILIRDKRFHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKKNILIRDKRFHYDR 600
QY 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISLRDQGTTFIDFKYNDKPLIYISNPNKYKNVAVTKENTIINPSNGDTSTNG 720
DB 661 DMLNISLRDQGTTFIDFKYNDKPLIYISNPNKYKNVAVTKENTIINPSNGDTSTNG 720
QY 721 IKKIL---IFSKKG 731
DB 721 IKKILKKVILGKG 734
```

```
RESULT 5
US-08-082-849B-12
; Sequence 12, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
```

```
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,849B
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,601
; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-161-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-082-849B-12
```

```
Query Match      98.7%; Score 3722.5; DB 1; Length 903;
Best Local Similarity 99.0%; Pred. No. 2.3e-253;
Matches 727; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 EVKQENRLNESSSQGLLYGFSDLNFOAPMVVTSSTTGLSPSSSELENIPSENOYF 60
DB 1 EVKQENRLNESSSQGLLYGFSDLNFOAPMVVTSSTTGLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTASDPSDFEKT 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTASDPSDFEKT 180
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKMWSTASDPSDFEKT 240
DB 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKMWSTASDPSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNYLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYNYLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWNYNQFLKTKQLRLDTQVYGNATYNFENGVRVDTGNSWSEV 480
```

Db 421 LNAQDDFSPTITMNNYQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNNWSEV 480
Qy 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
Qy 541 QYQKDIETEDFNDOOTSONIKQLAELNATNIYVLDKIKLNKQMLIRDKRFHYDR 600
Db 541 QYQKDIETEDFNDOOTSONIKQLAELNATNIYVLDKIKLNKQMLIRDKRFHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 661 DMLNISSLRODGKTFIDFKKYNKPLIYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGKTFIDFKKYNKPLIYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKIL---IFSKKG 731
Db 721 IKKILKVVLGKKG 734

RESULT 6

PCT-US94-01624-12
; Sequence 12, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-12

Query Match 98.7%; Score 3722.5; DB 5; Length 903;

Best Local Similarity 99.0%; Pred. No. 2.3e-253;

Matches 727; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

Qy 1 EVKQENRLNESSSQGLLYFSDLNFAQPMVTSSTTGDLSIPSSSELENIPSENQYF 60

Db 1 EVKQENRLNESSSQGLLYFSDLNFAQPMVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWGSFIVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIOY 120
Db 61 QSAIWGSFIVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIOY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVILSSDNQLPELKQKSSNSRKRKSTAGTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVILSSDNQLPELKQKSSNSRKRKSTAGTVPDRN 180
Qy 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPYSDPEKVT 240
Db 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPYSDPEKVT 240
Qy 241 GRIDKNVSPARPLVAAYPIVHVDMENILSKNEDQSTONTDSETRTISKNTSTSRHTT 300
Db 241 GRIDKNVSPARPLVAAYPIVHVDMENILSKNEDQSTONTDSETRTISKNTSTSRHTT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNNWSEV 480
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNNWSEV 480
Qy 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
Qy 541 QYQKDIETEDFNDOOTSONIKQLAELNATNIYVLDKIKLNKQMLIRDKRFHYDR 600
Db 541 QYQKDIETEDFNDOOTSONIKQLAELNATNIYVLDKIKLNKQMLIRDKRFHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 661 DMLNISSLRODGKTFIDFKKYNKPLIYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGKTFIDFKKYNKPLIYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKIL---IFSKKG 731
Db 721 IKKILKVVLGKKG 734

RESULT 7

US-08-082-849B-31
; Sequence 31, Application US/08082849B
; Patent No. 5677274

GENERAL INFORMATION:

APPLICANT: Leppla, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
TITLE OF INVENTION: Related Methods
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-31

Query Match          95.1%; Score 3590; DB 1; Length 719;
Best Local Similarity 95.4%; Pred. No. 3.5e-244;
Matches 705; Conservative 4; Mismatches 6; Indels 24; Gaps 2;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDYFTFATSADNHVTMWDDQEVINKASNKNIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYFTFATSADNHVTMWDDQEVINKASNKNIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS---RKGSTSAQPTVP 176
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS---RKGSTSAQPTVP 180
QY 177 DRNDGIPDSLEVEGYTVDNVKNKTFILSPWISNIHEKGLTKYKSSPEKWTASDPSDF 236
DB 181 DRNDGIPDSLEVEGYTVDNVKNKTFILSPWISNIHEKGLTKYKSSPEKWTASDPSDF 240
QY 237 EKVTRIDKNVSPARPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTS 296
DB 241 EKVTRIDKNVSPARPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTS 300
QY 297 RTHTSEVGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTAD 356
DB 301 RTHTSEVGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTAD 360
QY 357 TARLNANRYNTGTATYVNLPTTSLVGLKNOTLATIKAKENQLSOILAPNNYPSKML 416
DB 361 TARLNANRYNTGTATYVNLPTTSLVGLKNOTLATIKAKENQLSOILAPNNYPSKML 420
QY 417 APIALNAQKDFSPSTFITMYNQFLELEKTKQLRLDITDQVYGNATYNFENGRVVDTSN 476
DB 421 APIALNAQKDFSPSTFITMYN-----YGNATYNFENGRVVDTSN 460
QY 477 WSEVLPQIOETATIIIFNGKDLNIVERRIAANVPSDPLETTKPDNLTKEALKIATGFNEP 536
DB 461 WSEVLPQIOETATIIIFNGKDLNIVERRIAANVPSDPLETTKPDNLTKEALKIATGFNEP 520
QY 537 NGNLQYQGGKOITFDFNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKNLILIRDKRF 596
DB 521 NGNLQYQGGKOITFDFNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKNLILIRDKRF 580
QY 597 HYDRNNAVGADESVVKEAHEVINSSTEGLLNIDKIDKILSGYIVIEDTEGLKEVI 656
DB 581 HYDRNNAVGADESVVKEAHEVINSSTEGLLNIDKIDKILSGYIVIEDTEGLKEVI 640

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-31

Query Match          95.1%; Score 3590; DB 5; Length 719;
Best Local Similarity 95.4%; Pred. No. 3.5e-244;
Matches 705; Conservative 4; Mismatches 6; Indels 24; Gaps 2;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDYFTFATSADNHVTMWDDQEVINKASNKNIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYFTFATSADNHVTMWDDQEVINKASNKNIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS---RKGSTSAQPTVP 176
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS---RKGSTSAQPTVP 180
QY 177 DRNDGIPDSLEVEGYTVDNVKNKTFILSPWISNIHEKGLTKYKSSPEKWTASDPSDF 236
DB 181 DRNDGIPDSLEVEGYTVDNVKNKTFILSPWISNIHEKGLTKYKSSPEKWTASDPSDF 240
QY 237 EKVTRIDKNVSPARPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTS 296
DB 241 EKVTRIDKNVSPARPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTS 300
QY 297 RTHTSEVGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTAD 356
DB 301 RTHTSEVGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTAD 360
QY 357 TARLNANRYNTGTATYVNLPTTSLVGLKNOTLATIKAKENQLSOILAPNNYPSKML 416
DB 361 TARLNANRYNTGTATYVNLPTTSLVGLKNOTLATIKAKENQLSOILAPNNYPSKML 420
QY 417 APIALNAQKDFSPSTFITMYNQFLELEKTKQLRLDITDQVYGNATYNFENGRVVDTSN 476
DB 421 APIALNAQKDFSPSTFITMYN-----YGNATYNFENGRVVDTSN 460
QY 477 WSEVLPQIOETATIIIFNGKDLNIVERRIAANVPSDPLETTKPDNLTKEALKIATGFNEP 536
DB 461 WSEVLPQIOETATIIIFNGKDLNIVERRIAANVPSDPLETTKPDNLTKEALKIATGFNEP 520
QY 537 NGNLQYQGGKOITFDFNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKNLILIRDKRF 596
DB 521 NGNLQYQGGKOITFDFNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKNLILIRDKRF 580
QY 597 HYDRNNAVGADESVVKEAHEVINSSTEGLLNIDKIDKILSGYIVIEDTEGLKEVI 656
DB 581 HYDRNNAVGADESVVKEAHEVINSSTEGLLNIDKIDKILSGYIVIEDTEGLKEVI 640

657 NDYDMLNLISSLRQDGKTFIDFKYNDKLPYISNPNYKVNVAVTKENTINPSENGDT 716
641 NDYDMLNLISSLRQDGKTFIDFKYNDKLPYISNPNYKVNVAVTKENTINPSENGDT 700
QY 717 STNGIKKILIFSKGYEIG 735
DB 701 STNGIKKILIFSKGYEIG 719

PCT-US94-01624-31
Sequence 31, Application PC/TUS9401624
GENERAL INFORMATION:
APPLICANT: Leppia, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: ANTHEAX TOXIN FUSION PROTEINS AND
RELATED METHODS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
STREET: Steuart Street Tower, 20th Floor, One Market
STREET: Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01624
FILING DATE: June 25, 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-31
```

QY 237 EKVTGRIDKQVSPARHPPLVAAPIVHVDMENILSKNEDQSTONTDSETRTSKNTSTS 296
DB 241 EKVTGRIDKQVSPARHPPLVAAPIVHVDMENILSKNEDQSTONTDSETRTSKNTSTS 300
QY 297 RTHTSVHGNAEVAHSAFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTAD 356
DB 301 RTHTSVHGNAEVAHSAFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTAD 360
QY 357 TAPLNANIRVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQSLQILAPNNYPSKNL 416
DB 361 TAPLNANIRVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQSLQILAPNNYPSKNL 420
QY 417 APIALNAQDFSTPTIMNYNQFLEKTKQLRLDPTDQVYGNATYFNGRVRVDTGSN 476
DB 421 APIALNAQDFSTPTIMN-----YGNATYFNGRVRVDTGSN 460
QY 477 WSVLPQIOBTTHARILFNGKDLNVERRIAANVPSPLETTKPDMLTKEALKIARFENEP 536
DB 461 WSVLPQIOBTTHARILFNGKDLNVERRIAANVPSPLETTKPDMLTKEALKIARFENEP 520
QY 537 NGNLOYQGDITBFDNFQDQTSQNIKNQAEINATNIYVLDKIKNAKMNILIRDKRF 596
DB 521 NGNLOYQGDITBFDNFQDQTSQNIKNQAEINATNIYVLDKIKNAKMNILIRDKRF 580
QY 597 HYDRNNIAGADESVVKEAHREVINSSTGLLNIDKIRKILSGVIVIEDETEGLKEVI 556
DB 581 HYDRNNIAGADESVVKEAHREVINSSTGLLNIDKIRKILSGVIVIEDETEGLKEVI 640
QY 657 NDYDMNLNLSLQDQGTTFIDFKYNDKPLVYISPNYKVVAVYVTKNTIINPSENGDT 716
DB 641 NDYDMNLNLSLQDQGTTFIDFKYNDKPLVYISPNYKVVAVYVTKNTIINPSENGDT 700
QY 717 STNGIKKILFSGKGYEIG 735
DB 701 STNGIKKILFSGKGYEIG 719

RESULT 9
US-09-273-839A-8
; Sequence 8, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-273-839A-8

Query Match 33.8%; Score 1275; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 5.2e-82;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 466 ETTARIIFNGKDLNVERRIAANVPSPLETTKPDMLTKEALKIARFENEPNGLQYQK 545
DB 29 ETTARIIFNGKDLNVERRIAANVPSPLETTKPDMLTKEALKIARFENEPNGLQYQK 88
QY 546 DITEFDNFQDQTSQNIKNQAEINATNIYVLDKIKNAKMNILIRDKRFHYDRNNIAV 605
DB 89 DITEFDNFQDQTSQNIKNQAEINATNIYVLDKIKNAKMNILIRDKRFHYDRNNIAV 148
QY 606 GADES SVVKEAHREVINSSTGLLNIDKIRKILSGVIVIEDETEGLKEVINDRYDMLNI 665
DB 149 GADES SVVKEAHREVINSSTGLLNIDKIRKILSGVIVIEDETEGLKEVINDRYDMLNI 208

QY 666 SSLRQDQGTTFIDFKYNDKPLVYISPNYKVVAVYVTKNTIINPSENGDTSTNGIKKIL 725
DB 209 SSLRQDQGTTFIDFKYNDKPLVYISPNYKVVAVYVTKNTIINPSENGDTSTNGIKKIL 268
QY 726 IFSKKGYEIG 735
DB 269 IFSKKGYEIG 278

RESULT 10

US-08-960-780-32
; Sequence 32, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32608-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA-708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 177C8
US-08-960-780-32

Query Match 20.7%; Score 781; DB 3; Length 881;
Best Local Similarity 30.4%; Pred. No. 1.4e-46;
Matches 245; Conservative 134; Mismatches 298; Indels 130; Gaps 34;

QY 4 QENLLNSESQGLGYFSDLNFOAPVMVTSSTGDLSPSELEN--IPSENOYFQ 61
DB 39 QKQK---QKEMDRKGLGYFGKDF-SNLTFMFTPTDSTLIYDQQTANKLDDKQEQYQ 94
QY 62 SAINSGFVKVKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLEKGLYQIKIQYQ 121

Db 95 SIRWIGLQSKETGDTFNLSEDEQAIIEINGKIIISNGKKEQVHLEKGLVPIKIEYQ 154
Qy 122 RENPTKGLD-----FKLYWTDSONKKEVSSDNLQPLKQKSS-----N 162
Db 155 SD--TKFNIDSKTPEKELKFKIDSQOQVQOQDELNPFNFKESQEFLLAKPSKINLFT 212
Qy 163 SRKKESTAGPTVPRDNDGIPDSLEVEGYTVVKNKETFPLSPWISNIHEKGLTKYKSS 222
Db 213 QMKKEIDE-----TDTGDSIPDLWEENGYTI-----QNRIVAKWDDSL-ASKGYTKFVSN 264
Qy 223 PEKWTASDPYSDFEKTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNT 282
Db 265 PLESHTVGDPTDYDKAARDLDSNAKETFNPLVAAPSVNVSMKEVILSPENLS----- 320
Qy 283 DSETRTISKNTSTSRTHSEVHGNAEVSASFDDIGGSVAGFSNSNS---TVAIDHSL 339
Db 321 ----NSVESHSSTNWSY-----NTE-----GASVEAGIGPKGISFGVSUNYQHSET 363
Qy 340 LAGERTWAEITMG-----LNTADTARLNANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIK 395
Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANVRNNVGTGAIYVKPTTSFVL-NDTIAIT 420
Qy 396 AKENQLSQILAPNNYPSKNLAPIALNAQKDFSTPTIMVNOFLEKTKQLBLDQV 455
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHPITLNKKQVDNLLNPKMLETNQT 480
Qy 456 YGNIATYNPENGVRVDTGNSWSEVLPOIETATIIIFNGKDLNIVERRIAAVNSDPLE 515
Db 481 DG---VYK:KTHGNITVGEWNGVIOQIKAKTASIIVDDGE-RVAEKVAAKDYENPED 536
Qy 516 TTKPDMTLKAKIAP--GFNPENGLQYOGKDIETEDF--NFDOTSONKNQLAEL-- 569
Db 537 XT-PSLTLKDALKSLSPDEIKIEGLLYYKRPYESSVMTYLDENTAKETVKQLNDTGT 595
Qy 570 ---NATNIYVLKDKLNKAKMILLRDKRFHYDRNNIAGADESVVKEAHEVINSSTEG 626
Db 596 KFDVSHLYDV-----KLTFRMNTIK-LSILYDN---AESNDSIGKWTNIVSGGNG 647
Qy 627 -----LLANID-----KDIRKILGYIVEIDTE-----GLKE 654
Db 648 KQOYSSNNPDANLTNTLDAQELKNRDYVVISLYMKSEKNTQCEITIDGIEIYPTTKTVN 707
Qy 655 VINDRVDMLNI--SSLRQDKTPIDEKYNDKLPVWISPNKYVNAVYTKENTILNPSE 712
Db 708 VNKDNYKRLDIIAHNKSNPISIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763
Qy 713 NGDT-STNGIK---KILIFSKGYEIG 735
Db 764 IKQYISRYGKLEDGILIDKKGGIHYG 790

RESULT 11
US-09-073-898-32
; Sequence 32, Application US/09073898
; Patent No. 6242669
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacey
; TITLE OF INVENTION: No. 6242669a1 Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,898
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
APPLICATION DATA:
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 881 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: PSI177C8
US-09-073-898-32

Query Match 20.7%; Score 781; DB 3; Length 881;

Best Local Similarity 30.4%; Pred. No. 1.4e-46;

Matches 245; Conservative 134; Mismatches 298; Indels 130; Gaps 34;

Qy 4 QENRLNESSSQGLLYYFSDLNFOAPMVTSTTGDLSPSSLEN--TPSENQYFQ 61

Db 39 QKQK--QKMDRKGLLYYFKGKDF-SNLTMEFAPTRDSTLIYDQQTANKLDDKKQOEYQ 94

Qy 62 SAWSGFIKVKKDEYTFATSNHVTWVDDEVINKASNNKIRLEKRLYQIKIYQ 121

Db 95 SIRWIGLQSKETGDTFNLSEDEQAIIEINGKIIISNGKKEQVHLEKGLVPIKIEYQ 154

Qy 122 RENPTKGLD-----FKLYWTDSONKKEVSSDNLQPLKQKSS-----N 162

Db 155 SD--TKFNIDSKTPEKELKFKIDSQOQVQOQDELNPFNFKESQEFLLAKPSKINLFT 212

Qy 163 SRKKESTAGPTVPRDNDGIPDSLEVEGYTVVKNKETFPLSPWISNIHEKGLTKYKSS 222

Db 213 QMKKEIDE-----TDTGDSIPDLWEENGYTI-----QNRIVAKWDDSL-ASKGYTKFVSN 264

Qy 223 PEKWTASDPYSDFEKTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNT 282

Db 265 PLESHTVGDPTDYDKAARDLDSNAKETFNPLVAAPSVNVSMKEVILSPENLS----- 320

Qy 283 DSETRTISKNTSTSRTHSEVHGNAEVSASFDDIGGSVAGFSNSNS---TVAIDHSL 339

Db 321 ----NSVESHSSTNWSY-----NTE-----GASVEAGIGPKGISFGVSUNYQHSET 363

Qy 340 LAGERTWAEITMG-----LNTADTARLNANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIK 395

Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANVRNNVGTGAIYVKPTTSFVL-NDTIAIT 420

Qy 396 AKENQLSQILAPNNYPSKNLAPIALNAQKDFSTPTIMVNOFLEKTKQLBLDQV 455

Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHPITLNKKQVDNLLNPKMLETNQT 480

```

QY 456 YGNATYFNENGRVVDGTSNWSSEVLPOIQTETARIENGKDLNVERIAAIVNPSDPLE 515
Db 481 DG---VYKIDTHGNIVTGGWNGVIOQIKAKTASIIVDDGE-RVAEKVAAKOYENPED 536
QY 516 TTKPDMTLKEALKIAF--GFNEPNGNLOYGOKITEFDF--NFDQOTSQNIKNQIAEL-- 569
Db 537 KT-PSLTLDKALKUSYDPEIKIEGLLYKPKPIYESSVMYLDENTAKEVTKQLNDITG 595
QY 570 ---NATNIYTVLDKIKLNKAKNIIIRDKRPHYDRNNIAVGADESVVKEAHEVINSSTEG 626
Db 596 KFKDVSHLYDV---KLTPEKNNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647
QY 627 -----LLNID-----KDIRKILSGYIVEIEDTE-----GLKE 654
Db 648 KKOYSSNPNPDANLTNTDAQKLNKRDYVYISLYMKSEKNTQCEITIDGEIYPIITKTWN 707
QY 655 VINDRYDMLNI--SSLRODGKTFIDFKYNDKLPYISNPNYKNVYAVTKENTIIINSE 712
Db 708 VNKNYKRLDIIAHNKISNPISSIIHKT-NDEITLFWDDISI-TDVASIKPEN---LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKKGYEIG 735
Db 764 IKQIYSRYGKLEDGLIDKKGIIHYG 790

RESULT 12
US-09-307-106-8
; Sequence 8, Application US/09307106
; Patent No. 6603063
; GENERAL INFORMATION:
; APPLICANT: Peitelson, Jerald S.
; APPLICANT: Schnept, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacey
; TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/307,106
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/073,898
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355

```

```

; REFERENCE/DOCKET NUMBER: MA-708C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: PS177C8a
; US-09-307-106-8

```

```

Query Match 20.7%; Score 781; DB 4; Length 881;
Best Local Similarity 30.4%; Pred. No. 1.4e-46;
Matches 245; Conservative 134; Mismatches 298; Indels 130; Gaps 34;

QY 4 QENRLNESSSQGLGYFSDLNQAPMVTTSITGDISPSSLEN--IPSENQYFQ 61
Db 39 QKNQ---CKEMDRKGLLYGFKGDF-SNLTWFAPTDSTLIYDQOTANKLLKKQEQY 94
QY 62 SAWSGFIVKVKSDYEVTFATSDAHNVMTWVDDEVINKASNSKNIRLEKRLYQIKIQY 121
Db 95 SIRWIGLIQSKETGDTFNLSDEQALIEINGKIIISNKGKEQVHLEKGLVPKIEYQ 154
QY 122 RENPTEKGLD-----FKLYMTDSQNKKEVISSDNLQLPKOKSS-----N 162
Db 155 SD---TKFNIDSKTFKELKFKIDSQPOQVQODELENPEFNKESQEFLLAKSKINLFT 212
QY 163 SRKRSTSGAGTVPDRDNDGIPDSLEVEGYTVDKNKRFTLSPWISNIHEKGLTKYKSS 222
Db 213 QMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264
QY 223 PEKWTASDPYSDEPKYTKRIDKNVSPERARPLVAAVPIVHVDMENIILSKNEDQSTQNT 282
Db 265 PLESHTYGDPTDYDKAARDLDSNAKETFNPLVAFPPSVNVMKEVILSPNENLS----- 320
QY 283 DSETRISKNTSTSRTHSTSEVHGNAEVHASFPIGGSVSAGFSNSNSS-----TVAIDHSL 339
Db 321 ---NSVESHSSTWVSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 363
QY 340 LAGERTWAETWG-----INTADTARLNANIRYVNTGTAPIYNNVLTSLVLGKQCTLATIK 395
Db 364 VAQE--WGTSTGNTSQFTASAGYLNANVRNNVGTGAIYDVRPTTSFVL--NNDTIAIT 420
QY 396 AKENQLSQILAPNNYPSKNLAPIALNAQKDFSSPTITMNYNQFLEKTKQLRLDQV 455
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNSHPTILNKKQVDNLLNKKPMLETNQT 480
QY 456 YGNATYFNENGRVVDGTSNWSSEVLPOIQTETARIENGKDLNVERIAAIVNPSDPLE 515
Db 481 DG---VYKIDTHGNIVTGGWNGVIOQIKAKTASIIVDDGE-RVAEKVAAKOYENPED 536
QY 516 TTKPDMTLKEALKIAF--GFNEPNGNLOYGOKITEFDF--NFDQOTSQNIKNQIAEL-- 569
Db 537 KT-PSLTLDKALKUSYDPEIKIEGLLYKPKPIYESSVMYLDENTAKEVTKQLNDITG 595
QY 570 ---NATNIYTVLDKIKLNKAKNIIIRDKRPHYDRNNIAVGADESVVKEAHEVINSSTEG 626
Db 596 KFKDVSHLYDV---KLTPEKNNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647
QY 627 -----LLNID-----KDIRKILSGYIVEIEDTE-----GLKE 654
Db 648 KKOYSSNPNPDANLTNTDAQKLNKRDYVYISLYMKSEKNTQCEITIDGEIYPIITKTWN 707
QY 655 VINDRYDMLNI--SSLRODGKTFIDFKYNDKLPYISNPNYKNVYAVTKENTIIINSE 712
Db 708 VNKNYKRLDIIAHNKISNPISSIIHKT-NDEITLFWDDISI-TDVASIKPEN---LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKKGYEIG 735

```

Db 764 IKQYRYGKLEGGILIDKGGIHYG 790

RESULT 13

US-09-850-351A-32

Sequence 32, Application US/09850351A

Patent No. 6656908

GENERAL INFORMATION:

APPLICANT: Feltelson, Gerald S.

Schnepi, H. Ernest

Narva, Kenneth B.

Stockhoff, Brian A.

Schmeits, James

Loewer, David

Dullum, Charles Joseph

Muller-Cohn, Judy

Stamp, Lisa

Morrill, George

TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide Sequences Which Encode These Toxins

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606-6669

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/850,351A

FILING DATE: 07-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/073,898

FILING DATE: 06-MAY-1998

APPLICATION NUMBER: US 09/960,780

FILING DATE: 30-OCT-1997

APPLICATION NUMBER: US 60/029,848

FILING DATE: 30-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sanders, Jay M.

REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708CD1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 881 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: PS177C8

SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-09-850-351A-32

Query Match 20.7%; Score 781; DB 4; Length 881;

Best Local Similarity 30.4%; Pred. No. 1.4e-46;

Matches 245; Conservative 134; Mismatches 298; Indels 130; Gaps 34;

Qy 4 QENRLNESSSCGLLGYFSLNFQAPMVVTSSTGDLSPSSLEN--IPSENQYFQ 61

Db 39 QKQKQ---QKEMDRKGLLYGFKGDF-SNLTMPAPTRDSTLIYDQQTANKLLDKKQEQY 94

Qy 62 SATWSGFIKVKSDVEYFATSADNHVTMWDDQEVINKASNSKNKLEKGRLYQIKIQY 121

Db 95 SIRWIGLIQSKETGDTFNLSEDOAIIEINGKIISNKKEQVHVLEKGLVPIKIEYQ 154

Qy 122 RENPTEKGLD-----FKLYWTDSONKKEVIESDNLQPELKQKSS-----N 162

Db 155 SD--TKFNIDSKTFKELKLFKIDSONQPOQOQDELNPEFNKESQEBFLAKPSKINLFT 212

Qy 163 SRKRSTSGAPTVPDRDNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGKLTGYKSS 222

Db 213 QKMKREIDED---TDTGDSIPDLWEENGYTI-----QNRPIAVKWDSDL-ASKGYTFVSN 264

Qy 223 PEKWTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDOSTQMT 282

Db 265 PLESHTVGDPYTDYEKAARDLDSNAKETFNELVAAFPSVNVSMKVVILSPNENLS---- 320

Qy 283 DSETRISKNSTSRTHSEVHCGNAEVSASFDDIGGSVAGFSNNS--TVAIDHSLS 339

Db 321 ----NVESHSSINWSYT-----NTE-----GASVEAGIGPKGISFGVSINVYCHSET 363

Qy 340 LAGERTWAETMG---LNTADTARLNANTRYVNTGPAPYVNLPTTSLVLGKNQTLATIK 395

Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANRYVNVGTGAIYDVKPTTSFVL--NNDTIATIT 420

Qy 396 AKENQLSQILAPNNYPSKNLAPIALNAQKDSSTPITWYNOQFLEKTKOLRLDTDOV 455

Db 421 AKSNSTALNISPGESYKKGQNGIATSMDDFNHSBITLKKQVDNLLNKKPMMLSTNQ 480

Qy 456 YGNIATYFNGRVRVDTGSNMSEVLPIQIETTARIIFNGKDLNLVERIAAVNPSDPLE 515

Db 481 DG---VYKIDTRGNITVGTGEWNGVIQIKAKTASIIVDDGE--RVAEKVAAKDYENPED 536

Qy 516 TTXPDMTLKEALKIAP--GFNEPENGLOQCKDITEFDF--NPDQOTSONIKNQLAEL-- 569

Db 537 KT-PSLTLKDALKLSYDPEKEIEGLLYYKPKIYESSWVTLDETKAEVTKQLNDTTG 595

Qy 570 ---NATNIYTVLDKIKLNAMNIIIRDGRPHYDRNNIAVGADSVVKEAHRVINSSTG 626

Db 596 KFKDVSHLYDV---XLTPTKMNVTIK-LSILYDN---AESNDNSIGKWTNTIVSGNGNG 647

Qy 627 -----LLNID-----KDIRKLSGYIYEIEDTE-----GLXE 654

Db 648 KKOYSSNNPDANLTANTDAQEKLNKRDYIISLYMKSEKNTCEITDGEIYPIITKTWN 707

Qy 655 VINDRYDMLNI--SSLRQDGKTFIDFKKNDKLPYISNPYKRVYAVTKENTIINPSE 712

Db 708 VKNDYKELDIIAHNIKSNPISIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763

Qy 713 NGDT-STNGIK---KILIFSKKGYEIG 735

Db 764 IKQYRYGKLEGGILIDKGGIHYG 790

RESULT 14

US-08-471-033-5

Sequence 5, Application US/08471033

Patent No. 5770696

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W

APPLICANT: Kozziel, Michael G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J

APPLICANT: Carr, Brian

APPLICANT: Desai, Nalini M

APPLICANT: Kostichka, N. Kristy

APPLICANT: Duck, Nicholas B

APPLICANT: Estruch, Juan J

TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

516 TTKPDMTLKALKIAF--GFNEPNGNLQYQKDIETDF--NPDQOTSONIKNLQAL-- 569
540 KT-PSLTKALKSYDEIKIEGLLYKKNPIYESSVMTYLDENTAKEVTKQLNDTGG 598
570 ---NATNIYTVLDKIKLNAMKILIRDKRPHDRNNIAVGADESUVVKEAHRVINSSTEG 626
599 KFXDVSHLYDV---KLTPKMNVTIK-JSILYDN---AESNDNSIGKNTNTNIVSGNKG 650
627 -----LLLNID-----KDIRKILSGYIIEIDTE-----GLKE 654
651 KKOYSSNNPDANLTNTDAQKLNKRDYIISLYMKSEKNTQCEITIDGIEIPIITKTVN 710
655 VINDRYDMLN-----ISSLRDQGTFTDFKFKYNDKLPYISNPNYKVNVAVTKE 704
711 VNKDNYKRLDIIAHNKSNPISLSH-----IKTNDIEITLFWDDISI-TDVASIKPE 760
705 NTIINPSENGDT-STNGIK---KILIFSKGYEIG 735
761 N-LTDSEIKQIYSRYGKLEDCILIDKKGHIYG 793
RESULT 15
US-08-471-044-5
Sequence 5, Application US/08471044
Patent No. 5840868
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,044
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/463,483
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLV3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689

QY 516 TTKPDMTLKALKIAF--GFNEPNGNLQYQKDIETDF--NPDQOTSONIKNLQAL-- 569
Db 540 KT-PSLTKALKSYDEIKIEGLLYKKNPIYESSVMTYLDENTAKEVTKQLNDTGG 598
QY 570 ---NATNIYTVLDKIKLNAMKILIRDKRPHDRNNIAVGADESUVVKEAHRVINSSTEG 626
Db 599 KFXDVSHLYDV---KLTPKMNVTIK-JSILYDN---AESNDNSIGKNTNTNIVSGNKG 650
QY 627 -----LLLNID-----KDIRKILSGYIIEIDTE-----GLKE 654
Db 651 KKOYSSNNPDANLTNTDAQKLNKRDYIISLYMKSEKNTQCEITIDGIEIPIITKTVN 710
QY 655 VINDRYDMLN-----ISSLRDQGTFTDFKFKYNDKLPYISNPNYKVNVAVTKE 704
Db 711 VNKDNYKRLDIIAHNKSNPISLSH-----IKTNDIEITLFWDDISI-TDVASIKPE 760
QY 705 NTIINPSENGDT-STNGIK---KILIFSKGYEIG 735
Db 761 N-LTDSEIKQIYSRYGKLEDCILIDKKGHIYG 793

Query Match 20.6%; Score 779; DB 1; Length 884;
Best Local Similarity 30.3%; Pred. No. 1.9e-46;
Matches 247; Conservative 129; Mismatches 293; Indels 146; Gaps 34;
QY 4 QENRLNSESSESSQGLLYGFSDLNFPQPMVVTSTTGGDLSIPSELEN--IPSENQYFQ 61
Db 42 QXNQ---QKEMDRKGLLYGFYKGFQF-SNLTWFAETRDSTLIYDQOTANKLLDKKQEQY 97
QY 62 SAIWGFIKVKSDVEYFATSNADNHVTMMVDQEVINKASNNKTRLEKGLYQIKIYOY 121
Db 98 SIRWIGLIQSKETGDTFNLSEDEQAIIIEINGKIISNKGKQVHVHLEKGLVPIKIEY 157
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISSDNLQLPELKQSS-----N 162
Db 158 SD--TKFNIDSTFKELKLFKIDSONQOQOQOQDELNPFPNKKESQEFLLAKPSKINLFT 215
QY 163 SRKKSSTAGTVPDRNDGIPDSLEVEGYTVVKNKRTFLSPWTSNIHEKKGITKYSS 222
Db 216 QKMKREIDED---TDTGDSIPDLWEENGYTI---QNRIAVKWDDSL-ASKGYTKFYSN 267
QY 223 PEKWSSTASPYSDFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNT 282
Db 268 PLESHTVGPDIYDKAARDLDSNAKETFNELVAAPFVSVNVMSEKVLSPNENLS---- 323
QY 283 DSETRTSKNTSTSTHTSEVHGNAEVHASFFDIGGSVAGFSNENSS---TVAIDHSL 339
Db 324 ---NVESSHSTNNSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHSET 366
QY 340 LAGERTWATMG-----LNTADTARINANRYNTGTAPIYVLPPTSIVLGNQTLAIK 395
Db 367 VAGE--WGSTGNSTGNSTAGYLNANRYNVTGAIYDVKPTTSFVL--NNDTIATIT 423
QY 396 AKENQLSQILAFNNYPSKNLAFIALNAQKDFSSPTITMNNQFLELEKTKQLRLDTQV 455
Db 424 AKSNSTALNISPGESYPKKGQNGIALTSMDDNSHPITLKKQVDNLLNNKPMLETNQT 483
QY 456 YGNIATYNPENGVRVDTGNSNWEVLPOIQTETARIIFNGKOLNLVERRIAAVNSPDL 515
Db 484 DG---VYKIKDTHGNIVTGEWNGVIOQIKAKTASIIVDGGS-RVAEKRAVKADYENPED 539

QY 516 TTKPDMTLKALKIAF--GFNEPNGNLQYQKDIETDF--NPDQOTSONIKNLQAL-- 569
Db 540 KT-PSLTKALKSYDEIKIEGLLYKKNPIYESSVMTYLDENTAKEVTKQLNDTGG 598
QY 570 ---NATNIYTVLDKIKLNAMKILIRDKRPHDRNNIAVGADESUVVKEAHRVINSSTEG 626
Db 599 KFXDVSHLYDV---KLTPKMNVTIK-JSILYDN---AESNDNSIGKNTNTNIVSGNKG 650
QY 627 -----LLLNID-----KDIRKILSGYIIEIDTE-----GLKE 654
Db 651 KKOYSSNNPDANLTNTDAQKLNKRDYIISLYMKSEKNTQCEITIDGIEIPIITKTVN 710
QY 655 VINDRYDMLN-----ISSLRDQGTFTDFKFKYNDKLPYISNPNYKVNVAVTKE 704
Db 711 VNKDNYKRLDIIAHNKSNPISLSH-----IKTNDIEITLFWDDISI-TDVASIKPE 760
QY 705 NTIINPSENGDT-STNGIK---KILIFSKGYEIG 735
Db 761 N-LTDSEIKQIYSRYGKLEDCILIDKKGHIYG 793

US-08-471-033-5
REFERENCE/DOCKET NUMBER: P-40,403
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 884 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-471-044-5

QY 705 NTIINPSENGDT-STNGIK---KILIFSKKGYEIG 735 QY

0
0
0
2

0
0
A
0
.
0
H

.

0
E
H
0

Q
0
0

Db 61 QSAIWSGFIKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
QY 121 QRENPTKGLDPLKYWDTSQNKKEVITSSDNLOQLPELKQKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENPTKGLDPLKYWDTSQNKKEVITSSDNLOQLPELKQKSSNSRKRSTSGPTVPDRDN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNPNGNL 540
QY 541 QYQKDIETEDFDFNDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLILRDKRPHYDR 600
Db 541 QYQKDIETEDFDFNDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLILRDKRPHYDR 600
QY 601 NNTAVGADESVMKAEHREVNSTEGLLNIDKDKIRKLSGYIYVEIETEDGLKEVINDRY 660
Db 601 NNTAVGADESVMKAEHREVNSTEGLLNIDKDKIRKLSGYIYVEIETEDGLKEVINDRY 660
QY 661 DMLNISSLRQDKTFFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDKTFFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 2
US-09-848-909-1
; Sequence 1, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-1

Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFYFDLNFQAPMVVTSSTTGDLSIPSSSELENIIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYFYFDLNFQAPMVVTSSTTGDLSIPSSSELENIIPSENOYF 60
QY 61 QSAIWSGFIKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 61 QSAIWSGFIKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
QY 121 QRENPTKGLDPLKYWDTSQNKKEVITSSDNLOQLPELKQKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENPTKGLDPLKYWDTSQNKKEVITSSDNLOQLPELKQKSSNSRKRSTSGPTVPDRDN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNPNGNL 540
QY 541 QYQKDIETEDFDFNDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLILRDKRPHYDR 600
Db 541 QYQKDIETEDFDFNDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLILRDKRPHYDR 600
QY 601 NNTAVGADESVMKAEHREVNSTEGLLNIDKDKIRKLSGYIYVEIETEDGLKEVINDRY 660
Db 601 NNTAVGADESVMKAEHREVNSTEGLLNIDKDKIRKLSGYIYVEIETEDGLKEVINDRY 660
QY 661 DMLNISSLRQDKTFFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDKTFFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 3
US-09-848-909-2
; Sequence 2, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis

US-09-848-909-2

Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQNRLLNESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQNRLLNESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNQLPELKOKSSNSRKRSTSGPTVPPDRN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNQLPELKOKSSNSRKRSTSGPTVPPDRN 180
QY 181 QSAIWSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 181 QSAIWSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHAFDIDGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHAFDIDGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIRAKENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIRAKENQLSQIILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTTMYNQFLEKTKQLRLDQVYGNIAIYFNFGVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTTMYNQFLEKTKQLRLDQVYGNIAIYFNFGVRVDTGSNWSEV 480
QY 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
Db 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
QY 541 QYQKDIITFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAKQNTILIRDKRFHYDR 600
Db 541 QYQKDIITFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAKQNTILIRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGTFFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGTFFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 4
US-09-848-909-3
; Sequence 3, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04

US-09-848-909-3

Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQNRLLNESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQNRLLNESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNQLPELKOKSSNSRKRSTSGPTVPPDRN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNQLPELKOKSSNSRKRSTSGPTVPPDRN 180
QY 181 QSAIWSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 181 QSAIWSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHAFDIDGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHAFDIDGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIRAKENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIRAKENQLSQIILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTTMYNQFLEKTKQLRLDQVYGNIAIYFNFGVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTTMYNQFLEKTKQLRLDQVYGNIAIYFNFGVRVDTGSNWSEV 480
QY 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
Db 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
QY 541 QYQKDIITFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAKQNTILIRDKRFHYDR 600
Db 541 QYQKDIITFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAKQNTILIRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGTFFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGTFFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 5
US-09-848-909-4
; Sequence 4, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-09-848-909-4

Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLLYFSDLNFCAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSSQGLLYFSDLNFCAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKY 120
DB 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWSASDPSDFEKT 240
DB 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWSASDPSDFEKT 240
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480
QY 481 LPOIQUETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
DB 481 LPOIQUETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
QY 541 QYQKDIITEFDNFDQSTQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDQSTQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
QY 601 NNIAGADESVKAEHREVINSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660
DB 601 NNIAGADESVKAEHREVINSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660
QY 661 DMLNISLRQDGTFFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
DB 661 DMLNISLRQDGTFFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
QY 721 IKKILIPSKKGYEIG 735
DB 721 IKKILIPSKKGYEIG 735

RESULT 6
US-09-848-909-5

; Sequence 5, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-09-848-909-5

Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLLYFSDLNFCAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNESSSQGLLYFSDLNFCAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKY 120
DB 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWSASDPSDFEKT 240
DB 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWSASDPSDFEKT 240
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480
QY 481 LPOIQUETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
DB 481 LPOIQUETTARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
QY 541 QYQKDIITEFDNFDQSTQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDQSTQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
QY 601 NNIAGADESVKAEHREVINSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660
DB 601 NNIAGADESVKAEHREVINSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660
QY 661 DMLNISLRQDGTFFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
DB 661 DMLNISLRQDGTFFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
QY 721 IKKILIPSKKGYEIG 735

Db 721 IKKILIFSKGYEIG 735
|||||
RESULT 7
US-09-848-909-6
; Sequence 6, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-6
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENRLNESSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWGSFIVKVKSDSYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKYQ 120
Db 61 QSAIWGSFIVKVKSDSYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKYQ 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVTSNDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKEVTSNDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASDPYSDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASDPYSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSSTRTHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDWTLKEALKIAFGNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDWTLKEALKIAFGNEPENG 540
Qy 541 QYQKDIITFDNFDOQTQSNIKQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIITFDNFDOQTQSNIKQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
Qy 601 NNTAVGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVETEDTEGKEVINDRY 660
Db 601 NNTAVGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVETEDTEGKEVINDRY 660

Qy 661 DMLNISLRDGGKTPTDFKKYNDKPLXYISNPNYKVNVAVTYKNTIINPSENGDSTNG 720
Db 661 DMLNISLRDGGKTPTDFKKYNDKPLXYISNPNYKVNVAVTYKNTIINPSENGDSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735
RESULT 8
US-09-848-909-7
; Sequence 7, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-7
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENRLNESSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWGSFIVKVKSDSYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKYQ 120
Db 61 QSAIWGSFIVKVKSDSYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKYQ 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVTSNDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKEVTSNDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASDPYSDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASDPYSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSSTRTHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDWTLKEALKIAFGNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDWTLKEALKIAFGNEPENG 540
Qy 541 QYQKDIITFDNFDOQTQSNIKQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600

Db 541 QYQKDIETDFNFDOQTSONIKQLAELNATNIYTVLDKIKLAKNMILLIRKDFHYDR 600
QY 601 NNIAVGADESVEKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESVEKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRQDQKGTDFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDQKGTDFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
RESULT 9
US-09-848-909-8
; Sequence 8, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-8

Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDVTFATSAADNHVTMWVDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDVTFATSAADNHVTMWVDDQEVINKASNNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNNKRRKSTAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNNKRRKSTAGTPVDRDN 180
QY 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSADDPYSDPEKVT 240
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSADDPYSDPEKVT 240
QY 241 GRIDKNVSPARHPVLAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPARHPVLAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNNAEVAHSPFDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETGLTADTARL 360
Db 301 SEVHGNNAEVAHSPFDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETGLTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKAKENQLSQILAPNYYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKAKENQLSQILAPNYYPSKNLAPIA 420
QY 421 LNAQKDFSGSTPTMNYNFOFLEKTKQLRLDQVYGNIAATYNPENGRVDTGSGNWSEV 480
Db 421 LNAQKDFSGSTPTMNYNFOFLEKTKQLRLDQVYGNIAATYNPENGRVDTGSGNWSEV 480

QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLKEALKIAGFNEPENGNL 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLKEALKIAGFNEPENGNL 540
QY 541 QYQKDIETDFNFDOQTSONIKQLAELNATNIYTVLDKIKLAKNMILLIRKDFHYDR 600
Db 541 QYQKDIETDFNFDOQTSONIKQLAELNATNIYTVLDKIKLAKNMILLIRKDFHYDR 600
QY 601 NNIAVGADESVEKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESVEKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRQDQKGTDFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDQKGTDFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
RESULT 10
US-09-848-909-9
; Sequence 9, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-9
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDVTFATSAADNHVTMWVDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDVTFATSAADNHVTMWVDDQEVINKASNNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNNKRRKSTAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNNKRRKSTAGTPVDRDN 180
QY 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSADDPYSDPEKVT 240
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSADDPYSDPEKVT 240
QY 241 GRIDKNVSPARHPVLAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPARHPVLAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNNAEVAHSPFDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETGLTADTARL 360
Db 301 SEVHGNNAEVAHSPFDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETGLTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKAKENQLSQILAPNYYPSKNLAPIA 420


```
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTIMYVNOFLEKTKQLBLDTDOUVGNATYVNFENGRVVDGTGNSWSEV 480
Db 421 LNAQKDFSSPTIMYVNOFLEKTKQLBLDTDOUVGNATYVNFENGRVVDGTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Qy 541 QYQKDIITFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKAKNILLIRDKRFHYDR 600
Db 541 QYQKDIITFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKAKNILLIRDKRFHYDR 600
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
Qy 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Qy 721 IKKILIPSKGYEIG 735
Db 721 IKKILIPSKGYEIG 735
```

RESULT 11

US-09-848-909-10

; Sequence 10, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 736

; TYPE: PR

; ORGANISM: Bacillus anthracis

US-09-848-909-10

```
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSGDEYTFATSAONHVMTWVDQEVINKASNNKIRLEKRLYQIKIQY 120
Db 61 QSAIWSGFIKVKSGDEYTFATSAONHVMTWVDQEVINKASNNKIRLEKRLYQIKIQY 120
Qy 121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKRSSTAGTVPDRN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKRSSTAGTVPDRN 180
Qy 181 DGPDSLEVEGYTVDVNKNRTFLSPWISNIHEKKGLTKYKSSPEKSTADDPYDFEKT 240
Db 181 DGPDSLEVEGYTVDVNKNRTFLSPWISNIHEKKGLTKYKSSPEKSTADDPYDFEKT 240
Qy 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTQSTRTSTSTRTHT 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTQSTRTSTSTRTHT 300
```

```
Qy 301 SEVHGNAEVSFPDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Db 301 SEVHGNAEVSFPDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Qy 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTIMYVNOFLEKTKQLBLDTDOUVGNATYVNFENGRVVDGTGNSWSEV 480
Db 421 LNAQKDFSSPTIMYVNOFLEKTKQLBLDTDOUVGNATYVNFENGRVVDGTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Qy 541 QYQKDIITFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKAKNILLIRDKRFHYDR 600
Db 541 QYQKDIITFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKAKNILLIRDKRFHYDR 600
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIETDEGLKEVINDRY 660
Qy 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Qy 721 IKKILIPSKGYEIG 735
Db 721 IKKILIPSKGYEIG 735
```

RESULT 12

US-09-848-909-11

; Sequence 11, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11

; LENGTH: 736

; TYPE: PR

; ORGANISM: Bacillus anthracis

US-09-848-909-11

```
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSGDEYTFATSAONHVMTWVDQEVINKASNNKIRLEKRLYQIKIQY 120
Db 61 QSAIWSGFIKVKSGDEYTFATSAONHVMTWVDQEVINKASNNKIRLEKRLYQIKIQY 120
Qy 121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKRSSTAGTVPDRN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKRSSTAGTVPDRN 180
Qy 181 DGPDSLEVEGYTVDVNKNRTFLSPWISNIHEKKGLTKYKSSPEKSTADDPYDFEKT 240
```

```
Db 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYDFEYV 240
Qy 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVAHSPFDIGGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
Qy 421 LNAQDDFSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
Db 421 LNAQDDFSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAGFNEPENG 540
Qy 541 QYCGKDITEFDNFDOOTSQNIKNQLAELNATNITVLDKIKLNAKXNILLIRDKRPHYDR 600
Db 541 QYCGKDITEFDNFDOOTSQNIKNQLAELNATNITVLDKIKLNAKXNILLIRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 661 DMLNSSLRODGTFFIDFKYNDKPLYSNPNKYVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODGTFFIDFKYNDKPLYSNPNKYVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILFSKGYEIG 735
Db 721 IKKILFSKGYEIG 735
```

RESULT 13

```
US-09-848-909-12
; Sequence 12, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-12
```

```
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIVKYSDEYTFATSDAHNVMTWDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIVKYSDEYTFATSDAHNVMTWDDQEVINKASNNKIRLEKGRLYQIKIY 120
```

```
Qy 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQPLKQKSSNSRKRKSTASGTPVDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQPLKQKSSNSRKRKSTASGTPVDRDN 180
Qy 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYDFEYV 240
Db 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYDFEYV 240
Qy 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVAHSPFDIGGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
Qy 421 LNAQDDFSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
Db 421 LNAQDDFSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGVRVDTGNSWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAGFNEPENG 540
Qy 541 QYCGKDITEFDNFDOOTSQNIKNQLAELNATNITVLDKIKLNAKXNILLIRDKRPHYDR 600
Db 541 QYCGKDITEFDNFDOOTSQNIKNQLAELNATNITVLDKIKLNAKXNILLIRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 661 DMLNSSLRODGTFFIDFKYNDKPLYSNPNKYVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODGTFFIDFKYNDKPLYSNPNKYVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILFSKGYEIG 735
Db 721 IKKILFSKGYEIG 735
```

RESULT 14

```
US-09-848-909-13
; Sequence 13, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-13
```

```
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSPSSSELENIPSENQYF 60
```

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDQNKKEVYSSDNQLQLPELKQSSNSRKKSTASGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDQNKKEVYSSDNQLQLPELKQSSNSRKKSTASGPTVPDRDN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTYYKSSPKKSTASDPYDFPKVT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTYYKSSPKKSTASDPYDFPKVT 240
QY 241 GRIDKNVSPERARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
Db 241 GRIDKNVSPERARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
QY 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNIAATYFNENGRVRYVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNIAATYFNENGRVRYVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPENG 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRODQKTFIDFKKYNKDLPLYSNPNYKVNVAVTKENTIIINPSENGDSTNG 720
Db 661 DMLNISSLRODQKTFIDFKKYNKDLPLYSNPNYKVNVAVTKENTIIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 15
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-14

Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDQNKKEVYSSDNQLQLPELKQSSNSRKKSTASGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDQNKKEVYSSDNQLQLPELKQSSNSRKKSTASGPTVPDRDN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTYYKSSPKKSTASDPYDFPKVT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTYYKSSPKKSTASDPYDFPKVT 240
QY 241 GRIDKNVSPERARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
Db 241 GRIDKNVSPERARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
QY 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNIAATYFNENGRVRYVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNIAATYFNENGRVRYVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPENG 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRODQKTFIDFKKYNKDLPLYSNPNYKVNVAVTKENTIIINPSENGDSTNG 720
Db 661 DMLNISSLRODQKTFIDFKKYNKDLPLYSNPNYKVNVAVTKENTIIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

Search completed: May 3, 2004, 20:02:04
Job time : 53.6393 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 175.546 Seconds
(without alignments)
4086.665 Million cell updates/sec

Title: US-09-848-909A-8
Perfect score: 3773
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKKILIFSKKGYEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA Main:*

1:	/cgn2_6/prodata/2/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/prodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/prodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/prodata/2/paa/US08_COMB.pep.*
5:	/cgn2_6/prodata/2/paa/US09_COMB.pep.*
6:	/cgn2_6/prodata/2/paa/US10_COMB.pep.*
7:	/cgn2_6/prodata/2/paa/US11_COMB.pep.*
8:	/cgn2_6/prodata/2/paa/US12_COMB.pep.*
9:	/cgn2_6/prodata/2/paa/US13_COMB.pep.*
10:	/cgn2_6/prodata/2/paa/US14_COMB.pep.*
11:	/cgn2_6/prodata/2/paa/US15_COMB.pep.*
12:	/cgn2_6/prodata/2/paa/US16_COMB.pep.*
13:	/cgn2_6/prodata/2/paa/US17_COMB.pep.*
14:	/cgn2_6/prodata/2/paa/US18_COMB.pep.*
15:	/cgn2_6/prodata/2/paa/US19_COMB.pep.*
16:	/cgn2_6/prodata/2/paa/US20_COMB.pep.*
17:	/cgn2_6/prodata/2/paa/US21_COMB.pep.*
18:	/cgn2_6/prodata/2/paa/US22_COMB.pep.*
19:	/cgn2_6/prodata/2/paa/US23_COMB.pep.*
20:	/cgn2_6/prodata/2/paa/US24_COMB.pep.*
21:	/cgn2_6/prodata/2/paa/US25_COMB.pep.*
22:	/cgn2_6/prodata/2/paa/US26_COMB.pep.*
23:	/cgn2_6/prodata/2/paa/US27_COMB.pep.*
24:	/cgn2_6/prodata/2/paa/US28_COMB.pep.*
25:	/cgn2_6/prodata/2/paa/US29_COMB.pep.*
26:	/cgn2_6/prodata/2/paa/US30_COMB.pep.*
27:	/cgn2_6/prodata/2/paa/US31_COMB.pep.*
28:	/cgn2_6/prodata/2/paa/US32_COMB.pep.*
29:	/cgn2_6/prodata/2/paa/US33_COMB.pep.*
30:	/cgn2_6/prodata/2/paa/US34_COMB.pep.*
31:	/cgn2_6/prodata/2/paa/US35_COMB.pep.*
32:	/cgn2_6/prodata/2/paa/US36_COMB.pep.*
33:	/cgn2_6/prodata/2/paa/US37_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3773	100.0	735	DB 1	Length 735
2	3769	99.9	735	Pred. No. 5e-299	
3	3768	99.9	735	Mismatches 0	Indels 0 Gaps 0
4	3767	99.8	735	Conservative	

ALIGNMENTS

RESULT 1

PCT-US03-35733-8

Sequence 8, Application PC/TUS0335733

GENERAL INFORMATION:

APPLICANT: President and Fellows of Harvard College et al.

TITLE OF INVENTION: Compounds and Methods for the Treatment of Bacterial Infection

FILE REFERENCE: 00742/072003

CURRENT APPLICATION NUMBER: PCT/US03/35733

CURRENT FILING DATE: 2003-10-10

PRIOR APPLICATION NUMBER: US 60/424,987

PRIOR FILING DATE: 2002-11-08

NUMBER OF SEQ ID NOS: 38

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8

LENGTH: 735

TYPE: PRT

ORGANISM: Bacillus anthracis

PCT-US03-35733-8

Query Match 100.0%; Score 3773; DB 1; Length 735;

Best Local Similarity 100.0%; Pred. No. 5e-299;

Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60
Db 1 EVKQENRLNESSSSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTPATSDADNHVTWMDQEVINKASNSKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTPATSDADNHVTWMDQEVINKASNSKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDLQPELKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDLQPELKQKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDTPQVGNATYFNFNGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDTPQVGNATYFNFNGRVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERBRIAAVNPSPDLETKPDMTLKEALKIATFNGEPNGL 540
Db 481 LPOIQTETARIIFNGKDLNVERBRIAAVNPSPDLETKPDMTLKEALKIATFNGEPNGL 540
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATYIVVLDKIKLNAQNLILIRKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATYIVVLDKIKLNAQNLILIRKRFHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILFSKKGYEIG 735
Db 721 IKKILFSKKGYEIG 735

RESULT 2
US-09-848-909A-8
; Sequence 8, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis

US-09-848-909A-8
Query Match 100.0%; Score 3773; DB 23; Length 735;
Best Local Similarity 100.0%; Pred. No. 5e-299;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVKQENRLNESSSSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60
Db 1 EVKQENRLNESSSSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTPATSDADNHVTWMDQEVINKASNSKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTPATSDADNHVTWMDQEVINKASNSKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDLQPELKQKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDLQPELKQKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDTPQVGNATYFNFNGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDTPQVGNATYFNFNGRVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERBRIAAVNPSPDLETKPDMTLKEALKIATFNGEPNGL 540
Db 481 LPOIQTETARIIFNGKDLNVERBRIAAVNPSPDLETKPDMTLKEALKIATFNGEPNGL 540
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATYIVVLDKIKLNAQNLILIRKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATYIVVLDKIKLNAQNLILIRKRFHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILFSKKGYEIG 735
Db 721 IKKILFSKKGYEIG 735

RESULT 3
PCT-US03-35733-7
; Sequence 7, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-7

Query Match      99.9%; Score 3769; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.1e-298;
Matches 734; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKSTAGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKSTAGPTVPDRDN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSTPTTMNYNQFLELEKTKQLRLDQVYGNIAIYFNENGRVVDTSNWSFV 480
Db 421 LNAQKDFSTPTTMNYNQFLELEKTKQLRLDQVYGNIAIYFNENGRVVDTSNWSFV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 4
US-09-848-909A-7
; Sequence 7, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A

; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-7

Query Match      99.9%; Score 3769; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.1e-298;
Matches 734; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKSTAGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKSTAGPTVPDRDN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSTPTTMNYNQFLELEKTKQLRLDQVYGNIAIYFNENGRVVDTSNWSFV 480
Db 421 LNAQKDFSTPTTMNYNQFLELEKTKQLRLDQVYGNIAIYFNENGRVVDTSNWSFV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 5
PCT-US03-35733-6
; Sequence 6, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
```

; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-6

Query Match 99.9%; Score 3768; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.3e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60

Qy 61 QSAIWSGFIKVKSDSEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

Qy 121 QRENTEKGLDFKLYWTDSONKEVYSSDNLQPELKOKSNSRKRSTAGTPVDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVYSSDNLQPELKOKSNSRKRSTAGTPVDRDN 180

Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSDFEKT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSRTHT 300

Qy 301 SEVHGNAEYHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEYHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 361 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420

Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480

Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540

Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKMMNILLRDKRFFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKMMNILLRDKRFFHYDR 600

Qy 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDKIRKILSGYIVIEDETEGLKEVINDRY 660

Qy 661 DMLNISLRQDGKTFIDFKKYNKPLIYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISLRQDGKTFIDFKKYNKPLIYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720

Qy 721 IKKILIFSKKGYBIG 735
Db 721 IKKILIFSKKGYBIG 735

RESULT 6

US-09-848-909A-6
; Sequence 6, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-6

Query Match 99.9%; Score 3768; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.3e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60

Qy 61 QSAIWSGFIKVKSDSEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

Qy 121 QRENTEKGLDFKLYWTDSONKEVYSSDNLQPELKOKSNSRKRSTAGTPVDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVYSSDNLQPELKOKSNSRKRSTAGTPVDRDN 180

Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSDFEKT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSRTHT 300

Qy 301 SEVHGNAEYHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEYHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 361 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420

Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480

Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540

Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKMMNILLRDKRFFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKMMNILLRDKRFFHYDR 600

Qy 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDKIRKILSGYIVIEDETEGLKEVINDRY 660

Qy 661 DMLNISLRQDGKTFIDFKKYNKPLIYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISLRQDGKTFIDFKKYNKPLIYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720

Qy 721 IKKILIFSKKGYBIG 735

Db 721 IKKILFSGKGYEIG 735
RESULT 7
PCT-US03-35733-5
; Sequence 5, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-5
Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDSEYFATSAADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYFATSAADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVISSDNLPQELKQSSNSRKRSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKEVISSDNLPQELKQSSNSRKRSTAGTVPDRDN 180
Qy 181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWTASDPYDFEKT 240
Db 181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSTRITSKTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSTRITSKTSTSRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGSNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Qy 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFDFKYNKDKPLYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720

Db 661 DMLNISSLRQDGKTFDFKYNKDKPLYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILFSGKGYEIG 735
Db 721 IKKILFSGKGYEIG 735
RESULT 8
PCT-US03-35733-10
; Sequence 10, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-10
Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDSEYFATSAADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYFATSAADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVISSDNLPQELKQSSNSRKRSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKEVISSDNLPQELKQSSNSRKRSTAGTVPDRDN 180
Qy 181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWTASDPYDFEKT 240
Db 181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSTRITSKTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSTRITSKTSTSRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGSNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Qy 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660


```
Db 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 9
PCT-US03-35733-20
; Sequence 20, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-20

Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENOYF 60
Qy 61 QSAIWGFIKVKKSDYTFATSNADNHVTMWVDDQEVINKASNKIRLEKGLYQIKY 120
Db 61 QSAIWGFIKVKKSDYTFATSNADNHVTMWVDDQEVINKASNKIRLEKGLYQIKY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKSRKSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKSRKSTAGTVPDRDN 180
Qy 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSSTRTHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVLPPTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVLPPTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTIMNYNQFLEBKTQKLRDLTDQVYGNIAATYNFENGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTIMNYNQFLEBKTQKLRDLTDQVYGNIAATYNFENGRVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKOLNIVERRIAANVPSDPLETTKPDMTLKEALKAFGNPNGL 540
Db 481 LPOIQTETARIIFNGKOLNIVERRIAANVPSDPLETTKPDMTLKEALKAFGNPNGL 540
Qy 541 QYQKIDTEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKQNLIRDKRPHYDR 600
```

```
Db 541 QYQKIDTEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKQNLIRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 10
PCT-US03-35733-21
; Sequence 21, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-21

Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENOYF 60
Qy 61 QSAIWGFIKVKKSDYTFATSNADNHVTMWVDDQEVINKASNKIRLEKGLYQIKY 120
Db 61 QSAIWGFIKVKKSDYTFATSNADNHVTMWVDDQEVINKASNKIRLEKGLYQIKY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKSRKSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKSRKSTAGTVPDRDN 180
Qy 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSSTRTHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVLPPTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVLPPTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTIMNYNQFLEBKTQKLRDLTDQVYGNIAATYNFENGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTIMNYNQFLEBKTQKLRDLTDQVYGNIAATYNFENGRVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKOLNIVERRIAANVPSDPLETTKPDMTLKEALKAFGNPNGL 540
Db 481 LPOIQTETARIIFNGKOLNIVERRIAANVPSDPLETTKPDMTLKEALKAFGNPNGL 540
Qy 541 QYQKIDTEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKQNLIRDKRPHYDR 600
```

Db 481 LPOIQTETARIIFNGKDLNLVERRIAANPSDPLETTKPDMLKEALKIAFGNPNGL 540
QY 541 QYQKDIETEDFNFDQOOTSQNIKNQLAELNATNIYTVLDKIKLNAAWILIRDKRPHYDR 600
Db 541 QYQKDIETEDFNFDQOOTSQNIKNQLAELNATNIYTVLDKIKLNAAWILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRODQKTFIDFKKYNKPLIYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODQKTFIDFKKYNKPLIYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 11

US-09-791-537-43735
; Sequence 43735, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 43735
; LENGTH: 735
; TYPE: PRT
; ORGANISM: pdb 1ACC
US-09-791-537-43735

Query Match 99.8%; Score 3767; DB 22; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDPLKLYWTDSONKKEVISSDNLQPELKQKSSNRKKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDPLKLYWTDSONKKEVISSDNLQPELKQKSSNRKKRSTASGTPVDRDN 180
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH 300
QY 301 SEVHGNAEVHASFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Db 301 SEVHGNAEVHASFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
QY 361 NANIRYVNTGTAPIYVLPITSLVGLKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVLPITSLVGLKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDQVGNATYVNFENGRVAVDTGSKNWEV 480

Db 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDQVGNATYVNFENGRVAVDTGSKNWEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAANPSDPLETTKPDMLKEALKIAFGNPNGL 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAANPSDPLETTKPDMLKEALKIAFGNPNGL 540
QY 541 QYQKDIETEDFNFDQOOTSQNIKNQLAELNATNIYTVLDKIKLNAAWILIRDKRPHYDR 600
Db 541 QYQKDIETEDFNFDQOOTSQNIKNQLAELNATNIYTVLDKIKLNAAWILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRODQKTFIDFKKYNKPLIYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODQKTFIDFKKYNKPLIYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 12

US-09-848-909A-5
; Sequence 5, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-5

Query Match 99.8%; Score 3767; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDPLKLYWTDSONKKEVISSDNLQPELKQKSSNRKKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDPLKLYWTDSONKKEVISSDNLQPELKQKSSNRKKRSTASGTPVDRDN 180
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH 300
QY 301 SEVHGNAEVHASFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Db 301 SEVHGNAEVHASFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360

361 NANIRVNTGTAPIYNNVLTPTSLVLGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
Db
361 NANIRVNTGTAPIYNNVLTPTSLVLGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNNYQFLEKTKQLRDLTDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRDLTDQVYGNATYNFENGVRVDTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNVERRIIAAVNPSPLETTKPDMLTKEALKIAFGNPNGL 540
Db 481 LPQIQETTARIIFNGKDLNVERRIIAAVNPSPLETTKPDMLTKEALKIAFGNPNGL 540
Qy 541 QYQKDIITFDNFDOQTSONIKNQLAELNATNIYTLVDKIKLNAXNNILIRDKRPHYDR 600
Db 541 QYQKDIITFDNFDOQTSONIKNQLAELNATNIYTLVDKIKLNAXNNILIRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIPSKKGYEIG 735
Db 721 IKKILIPSKKGYEIG 735
RESULT 13
US-09-848-909A-10
; Sequence 10, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848, 909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201, 800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-10
Query Match 99.8%; Score 3767; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENLLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYF 60
Db 1 EVKQENLLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYF 60
Qy 61 QSAIWSGFIKVKSDYFTFATSADNHVTMVDQEVINKASNNKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYFTFATSADNHVTMVDQEVINKASNNKIRLEKRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTVPDRN 180
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTADPYSDFEKVT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTADPYSDFEKVT 240
Qy 241 GRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDOSTQNTDSETRTKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAPYPIVHVDMENIILSKNEDOSTQNTDSETRTKNSTSRHT 300

301 SEYHGNAEVHASFDDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEYHGNAEVHASFDDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRVNTGTAPIYNNVLTPTSLVLGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYNNVLTPTSLVLGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNNYQFLEKTKQLRDLTDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRDLTDQVYGNATYNFENGVRVDTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNVERRIIAAVNPSPLETTKPDMLTKEALKIAFGNPNGL 540
Db 481 LPQIQETTARIIFNGKDLNVERRIIAAVNPSPLETTKPDMLTKEALKIAFGNPNGL 540
Qy 541 QYQKDIITFDNFDOQTSONIKNQLAELNATNIYTLVDKIKLNAXNNILIRDKRPHYDR 600
Db 541 QYQKDIITFDNFDOQTSONIKNQLAELNATNIYTLVDKIKLNAXNNILIRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Qy 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIPSKKGYEIG 735
Db 721 IKKILIPSKKGYEIG 735
RESULT 14
US-09-848-909A-20
; Sequence 20, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848, 909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201, 800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-20
Query Match 99.8%; Score 3767; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENLLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYF 60
Db 1 EVKQENLLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYF 60
Qy 61 QSAIWSGFIKVKSDYFTFATSADNHVTMVDQEVINKASNNKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYFTFATSADNHVTMVDQEVINKASNNKIRLEKRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTVPDRN 180
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTADPYSDFEKVT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTADPYSDFEKVT 240

Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSOILLAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSOILLAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDQVYGNIAATYVNFENGVRVDTGNNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDQVYGNIAATYVNFENGVRVDTGNNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGFNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGFNEPNGNL 540
Qy 541 QYQCKDITEFDFNFQQTSONIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Db 541 QYQCKDITEFDFNFQQTSONIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATYVNTKNTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATYVNTKNTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 15
US-09-848-909A-21
; Sequence 21, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002 US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-21

Query Match 99.8%; Score 3767; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENRLNESESSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESESSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFTKVKKSDYEYFATADNHVTWVDDQEVINKASNSNKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFTKVKKSDYEYFATADNHVTWVDDQEVINKASNSNKIRLEKRLYQIKIY 120
Qy 121 ORENPTKEGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKRSAGTVPDRDN 180

Db 121 ORENPTKEGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKRSAGTVPDRDN 180
Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKVT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSOILLAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSOILLAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDQVYGNIAATYVNFENGVRVDTGNNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDQVYGNIAATYVNFENGVRVDTGNNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGFNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGFNEPNGNL 540
Qy 541 QYQCKDITEFDFNFQQTSONIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Db 541 QYQCKDITEFDFNFQQTSONIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATYVNTKNTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATYVNTKNTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

Search completed: May 3, 2004, 19:57:37
Job time : 190.546 secs

OM protein - protein search, using sw model
Run on: May 3, 2004, 19:30:47 ; Search time 13.0855 Seconds
(without alignments)
5403.004 Million cell updates/sec

Title: US-09-848-909A-8
Perfect score: 3773
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKILFFSKGYEIG 735
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: piri: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3767	99.8	764	2 I39934	protective antigen
2	877	23.2	875	2 I40862	lota toxin compo
3	235.5	6.2	192	2 I39933	cryptic protein -
4	235.5	6.2	204	2 G59104	hypothetical prote
5	205	5.4	4688	2 F82885	hypothetical prote
6	187.5	5.0	2401	2 T28576	rhopty protein -
7	185.5	4.9	1639	2 S05503	major merozoite su
8	185	4.3	2259	2 T28577	rhopty protein -
9	184.5	4.9	2529	2 B64535	toxin-like outer m
10	183.5	4.9	1125	2 E90598	membrane nuclease,
11	183	4.9	4152	2 T31102	filamentous hemag
12	182	4.8	6713	2 B89921	hypothetical prote
13	180	4.8	1193	2 S68218	botulinum neurotox
14	179	4.7	1227	2 C97033	uncharacterized pr
15	178.5	4.7	1072	2 A86827	hypothetical prote
16	178	4.7	1302	1 JC6009	surface-located me
17	178	4.7	4919	2 T31105	hypothetical prote
18	177.5	4.7	821	2 S67087	hypothetical prote
19	177	4.7	752	2 G90599	hypothetical prote
20	177	4.7	1635	2 A10452	hemolysin [impor
21	175	4.6	1658	2 S55101	hypothetical prote
22	174.5	4.6	1365	2 T30822	impli protein - Myc
23	173	4.6	5005	2 F82884	hypothetical prote
24	172.5	4.6	2399	2 H71879	toxin-like outer m
25	172	4.6	1939	2 T18372	repeat organellar
26	171	4.5	2178	2 S55805	alpha-toxin - Clos
27	170	4.5	1631	1 SAZOK1	major merozoite su
28	170	4.5	1837	2 T41023	probable nuclear p
29	169.5	4.5	1230	2 S56850	SWC1 protein homol

30 169.5 4.5 2340 2 B71704 cell surface antig
31 168 4.5 1315 2 T28679 fibrinogen-binding
32 168 4.5 1957 2 T38077 hypothetical coile
33 166.5 4.4 1051 2 T18351 impli protein - Myc
34 166.5 4.4 1802 2 S52611 Tys protein - Yeas
35 166.5 4.4 2334 2 S32920 cell wall-associat
36 166 4.4 3724 2 T18427 hypothetical prote
37 165.5 4.4 786 2 T18469 hypothetical prote
38 165.5 4.4 979 2 JQ0894 P115 protein - Myc
39 165.5 4.4 1115 2 T41342 probable coiled-co
40 165 4.4 1308 2 E71622 probable membrane
41 165 4.4 3216 2 C90538 hypothetical prote
42 164 4.3 1465 2 S31262 Tys protein - Yeas
43 164 4.3 1516 2 E71619 RAD2 endonuclease
44 164 4.3 1803 2 S56894 Tys protein - Yeas
45 163.5 4.3 624 2 PC6003 surface membrane p

ALIGNMENTS

RESULT 1
I39934
Protective antigen precursor - Bacillus anthracis plasmid
C:Species: Bacillus anthracis
C:Date: 19-Jul-1996 #sequence,revision 19-Jul-1996 #text_change 01-Dec-2000
C:Accession: I39934; S69160; F59104
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr
A:Reference number: I39933; MUID:89172073; PMID:3148491
A:Accession: I39934
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-764 <RES>
A:Cross-references: GB:M22589; NID:g143280; PIDN:AAA22637.1; PID:g143282
R:Friedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klimpel, K.R.; Birch, N.P.; Loh, Y.P.
Arch. Biochem. Biophys. 316, 5-13, 1995
A:Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SF
A:Reference number: S69160; MUID:95142670; PMID:7840657
A:Accession: S69160
A:Molecule type: protein
A:Residues: 197-202 <FRI>
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koshle
J. Bacteriol. 181, 6509-6515, 1999
A:Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbor
A:Reference number: A59091; MUID:99445483; PMID:10515943
A:Accession: F59104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-313 'O', 315-764 <OKI>
A:Cross-references: GB:AF065404; NID:g4894216; PIDN:AAD32414.1; PID:g4894326
A:Experimental source: strain Sterne
A:Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid
C:Genetics:
A:Gene: pXOI-110
A:Genome: plasmid
C:Function:
A:Description: three component exotoxin; protective antigen binds to receptors on the s
y active components edema factor or lethal factor; the complex is internalized by recep
C:Keywords: exotoxin
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-196/Domain: propeptide #status predicted <PRO>
F:197-202/Product: protective antigen #status experimental <MAT>

Query Match 99.8%; Score 3767; DB 2; Length 764;

Best Local Similarity 99.9%; Pred. No. 1.7e-182;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYYFSDLNFOAPMVVVTSTTGDLSIPSELENIPSENYQYF 60
DB 30 EVKQENRLNSESQGLLYYFSDLNFOAPMVVVTSTTGDLSIPSELENIPSENYQYF 89

QY 61 QSAIWSGFTKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKRLYOIKIQ 120
 DB 90 QSAIWSGFTKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKRLYOIKIQ 149
 QY 121 QRENPTKGLDFKLYWTDSONKKEVYSSDNLQPELKQKSSNSRKRSTAGTPVDRDN 180
 DB 150 QRENPTKGLDFKLYWTDSONKKEVYSSDNLQPELKQKSSNSRKRSTAGTPVDRDN 209
 QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWSTASDPYSDFEKT 240
 DB 210 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWSTASDPYSDFEKT 269
 QY 241 GRIDKNSPEARHPLVAAPYVHVDMENIILSKNEDQSTONTSDSTRISKNSTSRHT 300
 DB 270 GRIDKNSPEARHPLVAAPYVHVDMENIILSKNEDQSTONTSDSTRISKNSTSRHT 329
 QY 301 SEVHGNAEYHAGFFDGGSVSAGFSNSNSTVAIDHSLAGERTWAEETMGLNTADTARL 360
 DB 330 SEVHGNAEYHAGFFDGGSVSAGFSNSNSTVAIDHSLAGERTWAEETMGLNTADTARL 389
 QY 361 NANIRVNTGTAPIYVNLPTSLVLGKNQTLATIKAKENQLSQIILAPNNYPSKNLAPTA 420
 DB 390 NANIRVNTGTAPIYVNLPTSLVLGKNQTLATIKAKENQLSQIILAPNNYPSKNLAPTA 449
 QY 421 LNAQKDFSSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
 DB 450 LNAQKDFSSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 509
 QY 481 LPOIQTETARIIFNGKDLNIVERIAAVNPSPLETTKPDMLTKEALKIATFNGPGL 540
 DB 510 LPOIQTETARIIFNGKDLNIVERIAAVNPSPLETTKPDMLTKEALKIATFNGPGL 569
 QY 541 QYQKDIETDFDFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
 DB 570 QYQKDIETDFDFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 629
 QY 601 NNTAVGADESVEKAEHREVNSTEGLLNIDKIDIKLSGYIVETEDTEGLKEVINDRY 660
 DB 630 NNTAVGADESVEKAEHREVNSTEGLLNIDKIDIKLSGYIVETEDTEGLKEVINDRY 689
 QY 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNKYVNVYAVTKNTIINPSENGDTSTNG 720
 DB 690 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNKYVNVYAVTKNTIINPSENGDTSTNG 749
 QY 721 IKKILIFSKGYEIG 735
 DB 750 IKKILIFSKGYEIG 764

RESULT 2
 I40862
 Iota toxin component Ib - Clostridium perfringens
 C:Species: Clostridium perfringens
 C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Oct-1999
 C:Accession: I40862; S42774
 R:Perelle, S.; Gilbert, M.; Boquet, P.; Popoff, M.R.
 Infect. Immun. 61, 5147-5156, 1993
 A:Title: Characterization of Clostridium perfringens Iota-toxin genes and expression in
 A:Reference number: I40861; MUID:94041637; PMID:8225592
 A:Accession: I40862
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-975 <RES>
 A:Cross-references: EMBL:X73562; NID:g929031; PIDN:CAA51960.1; PID:g414655

Query Match 23.2%; Score 877; DB 2; Length 875;
 Best Local Similarity 31.2%; Pred. No. 1.1e-36;
 Matches 253; Conservative 131; Mismatches 277; Indels 150; Gaps 31;

QY 1 EVKQENRLNSESSESSQGLGYTSDNLQFAPMVVTSSTGDISIPSELENIPSE-NQY 59
 DB 35 DTNQKEETNTLSSNGLMGYYFADEBFKDLLEWAPIKNGDLKAFBEKKYDKLLTENDSS 94

QY 60 FQSAIWSGFTKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKRLYOIKIQ 119
 DB 95 IKSIRWTGRIIPSEDEGEYILSDTR-NDVLMQINAKGDIK-...TLKVMKKQCAYNIRIE 150
 QY 120 YQREN-----PTEKGLDFKLYWTDSONKKEVYSSDNLQPELKQKSSNSRKRST 170
 DB 151 IQDKNLGSDNLGVP-----KLYW-ELNGKNTVIPENLFFRDYDKIDEND----- 195
 QY 171 AGTVP-----DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNTH 211
 DB 196 --PFIPNNNFVFRFSAAWEDDLTDNDNIPDAYEKNGYII-----KDSIAVKNWDSFA 249
 QY 212 EKEGLTKYKSPKSWSTASDPYSDFEKTGRIDKNSPEARHPLVAAPYVHVDMENIIL 271
 DB 250 E-QYKYYSSYLESTAGDPYTDYQKASGSDKAIKLEARDPLVAAPYVGVGMENLII 308
 QY 272 SKNEDQSTQNTDSETRISKNSTSTHSEVHGNAEYHAGFFDGGSVSAGFSNSNST 331
 DB 309 STNEHASS-----DQKTVSRATNSKTDAITV-----GVSTISAGVQNGFTGN 351
 QY 332 VAIDHS-----LSLAGERTWAEETMGLNTADTARLNANIRVNTGTAPIYVNLPTSL 383
 DB 352 ITTSYSHTTDNTAVQDSNGESNWTGLSINKGESAVINANVRVYNTGTAPMYKVTPTNL 411
 QY 384 VLGNQTLATIKAKENQLSQIILAPNNYPSKNLAPTAALNAQKDFSSPTIMYNNQFLE 443
 DB 412 VL-DGETLATIKADQNGQIGNNLSFNETYPKKGLSPALNTMDQFNARLIPINDYDQKKLD 470
 QY 444 KTKQLRLDQVYGNATYNFENGVRVDTGNSWSEVLPQIQTETARIIFNGKDLNIVER 503
 DB 471 SGQIKLETTQVSGNYCTKN-SQOI-ITEGNSKSNVISQIDSVSASIILD-TGSQTFER 527
 QY 504 RIAAVNSDPLETTKPDMLTKEALKIATFNGPENGNYOY-OGKDIET--FDFNPDQOTSQ 560
 DB 528 RVAAKEQGNPEKPT-PEITIGEAIKKAFSATK-NGELLYFNGIPIDESCVELIFDDNTSE 585
 QY 561 NIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKR--HYDR-NNTAVGADESVEKAEH 617
 DB 586 IIKQKLYLDDKLYNV---KLERGNVILIKVPSYFTNFEYNNFP--ASWSNIDTFNQ 639
 QY 618 EVINSSTEGT-----LLNIDKIDIKLSGY-----IVIEITEGLKEVIND 658
 DB 640 DGLQSVANKLSGETKIIIPMSKLPKRYVFSGYKSDPSTNSITVNIKSKEQKTDYLV 699
 QY 659 RYDMLNTS-----SLRQDGKTFIDFKYND--KLPLYISNPNKYV----- 696
 DB 700 EKDYTKSYEFETGKDSSEIETLTSSGVIFLDNLSITELNSTPEILKEPEIKVPSDQE 759
 QY 697 -----NVYAVTKNTIINPSENGDTSTNGI 721
 DB 760 ILDAHKNKYADIKLDT-----NTGNTYIDGI 785

RESULT 3
 I39933
 Cryptic protein - Bacillus anthracis
 C:Species: Bacillus anthracis
 C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999
 C:Accession: I39933
 R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
 Gene 69, 287-300, 1986
 A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr
 A:Reference number: I39933; MUID:89172073; PMID:3148491
 A:Accession: I39933
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-192 <RES>
 A:Cross-references: GB:M22589; NID:g143280; PIDN:AAA22636.1; PID:g143281

Query Match 6.2%; Score 235.5; DB 2; Length 192;
 Best Local Similarity 34.6%; Pred. No. 2.6e-05;
 Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNILIRKDFHYDRNNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEI 646
Db 1 MNILVRDP-YHYDNGNIVGVDDSYLKAYKQILNWSGVSINLDEVDNQALSGYMLQI 59
QY 647 EDTE-----GLKEVINRDMYKMLISSLRQDKTFIDFKKXNDKLPYISNPN 693
Db 60 KKPSNHLTNSPVTTITLAGKDSGVGLYRVL-----DQAGFLDNKFDENRSLV-DPG 112
QY 694 YKVNVAVTNKEN-TIINPSENGDTSTNGIKKILIFSKGGEI 734
Db 113 DDVYVAVTKEDFNATVDRNGNIA-NKLKNTLVLSGKIKEI 153

RESULT 4
G59104
hypothetical protein pXOI-111 - Bacillus anthracis virulence plasmid pXOI
C:Species: Bacillus anthracis
C>Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-May-2000
C:Accession: G59104
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler, J. Bacteriol. 181, 6509-6515, 1999
A:Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbored by B. anthracis
A:Reference number: A59091; MUID:99445483; PMID:10515943
A:Accession: G59104
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <OKI>
A:Cross-references: GB:AF065404; NID:94894216; PIDN:AA032415.1; PID:94894327
A:Experimental source: strain Sterne
A>Note: similar to hypothetical protein in the protective antigen domain; ypa, plasmid pXOI-111
C:Genetics:
A:Gene: pXOI-111
A:Genome: plasmid

Query Match 6.2%; Score 235.5; DB 2; Length 204;
Best Local Similarity 34.6%; Pred. No. 2.9e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNILIRKDFHYDRNNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEI 646
Db 1 MNILVRDP-YHYDNGNIVGVDDSYLKAYKQILNWSGVSINLDEVDNQALSGYMLQI 59
QY 647 EDTE-----GLKEVINRDMYKMLISSLRQDKTFIDFKKXNDKLPYISNPN 693
Db 60 KKPSNHLTNSPVTTITLAGKDSGVGLYRVL-----DGTGFLDNKFDENRSLV-DPG 112
QY 694 YKVNVAVTNKEN-TIINPSENGDTSTNGIKKILIFSKGGEI 734
Db 113 DDVYVAVTKEDFNATVDRNGNIA-NKLKNTLVLSGKIKEI 153

RESULT 5
F82885
hypothetical protein U0482 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82885
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a minor variant
A:Reference number: A82870
A:Accession: F82885
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4688 <GLA>
A:Cross-references: GB:AE002145; GB:AF222894; NID:96899476; PIDN:AAF30894.1; GSFDB:GN001
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: U0482
A:Genetic code: SGC3

Query Match 5.4%; Score 205; DB 2; Length 4688;
Best Local Similarity 21.3%; Pred. No. 0.1;

Matches 174; Conservative 132; Mismatches 324; Indels 186; Gaps 40;
QY 19 LLGYFDFNFOAPMVVTSSTTGLSIPSELENIPSENQVFSQAIWSGFIKVKSDYET 78
Db 3699 LVDVYLD-NIHQNDITETKIFKHNV-SKEIENPGVTMISKHGWKSPDTDTTANFEK 3756
QY 79 FATSADNHTVMVDDDEVINKASNKIRLEKELYQIKIQYQRENPTKGLDFKLYWTD 138
Db 3757 IETQ-----DDNVLNNDATVFKDEHNHKKQIVRIKEN-----ND 3795
QY 139 SQNKKEVISDNLQLPKOKSN-----SRKKRSTSGPTVDPDR----- 179
Db 3796 WLKQGI---DNLN-PETKYKLENIELSKPLKTHHTNLSVINDKENISLITETGNPVLKV 3851
QY 180 ----NGIPDSLEVEGYTVQVKNKRFPLSPWISNIEHKKGLTKYKSPKSWSTASDPYSD 235
Db 3852 IQQNDTINDTQQTINVTLSGVNSK-YNGROIKVYKDNNNVIYESS---LITLQKGDND 3907
QY 236 FEKVTGRIDKNSPEARHPLVAAPVIVHVMENILSKN-EDQSTQNTDSET-RTISKNT 293
Db 3908 YQLLNLSNEN-----REYREKIEINHSNTNNFEDEKLGKNGVNFITQKNT 3957
QY 294 STSRHTS-EVHGNAEVAHAF-----FDIGGSVSAGFS-----NSNS---STVAID 335
Db 3958 TVQWNDSSATVGRGVNFNFKIKSBDKILENNQVVAWFAPKETTIRDTNTWLQYTRPLK 4017
QY 336 HSLSLAGERTWAETMGINT---ADTAELNANIRYVNTGAPVIVLPTTSLVLGKQOTLA 392
Db 4018 DVTSDFKEGWADLNSVNVKPEETTKLVKIQVKNPKAKNNINNSVILDTNISI 4077
QY 393 -----TIKAKENQLSQILAPNNYPSKNLAPALN-AQDFSSPTITWNY--NQFLELE 443
Db 4078 NSNYEFTTKVDHKLINITSSNNVNTNSQTINFTLSGVKKSVMGKKIKLSYKSNDSSES 4137
QY 444 KTKQLRLDTPQVYGNIAATYVFNENR-----VEVDGSGNWSVLPQIQTETARIIFNGKD 497
Db 4138 HTMEVLIESNKTQVNIILLNKLKRNRYTLLDVLKIDNNVSDPFKSGNLNPSITRTSA 4197
QY 498 LNLVERRIAAVNSDPLETT-----KPDMLKEA-----LKIAPGFNEPGLNQ 541
Db 4198 INVNLIEISNRASTNLKSTIIKINLNDPDLVLRDQDQATVYGNKKNQAMGFITVSGNIK 4257
QY 542 YQKWDITEFPNPDQTSQNIK-NQLAELNATNIIYVLDKIKLNKNNILIRDKRPHYD- 599
Db 4258 YLTATLVDLNFN-DKVNIVNISFNKPSIAAEN-----IGIDKNNII-----YNND 4304
QY 600 -----RNNTAVGA---DESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIVEIEDTE 650
Db 4305 IPKLEINNDIIVNGPINKKEIVVKNANQX--NNIDVDLGLQINPKIAHLA-FIAKPKSTN 4361
QY 651 GLKEVINRDMYKMLISSLRQDKTFIDF-----KKY----- 681
Db 4362 -----NDIITNVTNGSSLVNNDGKTSIRFTLNLKANKLYSLVDVYVYLVNNSNIVE 4415
QY 682 NDKLPYISNPNYKVNVAVTNKENIINPSENGDTS 717
Db 4416 SNKLP-KLNNINYQIK---INKSHITII--SKNGEWS 4445

RESULT 6
T28676
rhoptry protein - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28676; A45521
R:Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem. Parasitol. 76, 329-332, 1996
A:Title: Comparison of two members of a multigene family coding for high-molecular mass
A:Reference number: 220507; MUID:97077455; PMID:8920022
A:Accession: T28676
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2401 <SIN>

A;Cross-references: EMBL:U36927; NID:gl041784; PID:gl041785; PIDN:AAB41263.1
R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A;Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co
A;Reference number: A45521; MUID:91101660; PMID:2270106
A;Accession: A45521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 2260-2401 <KEP>
A;Cross-references: GB:M34281

Query Match 5.0%; Score 187.5; DB 2; Length 2401;
Best Local Similarity 21.7%; Pred. No. 0.29;
Matches 182; Conservative 118; Mismatches 317; Indels 221; Gaps 43;

QY 1 EVKQENRLNESESSQGLLYYFSDLN-----FOAPMVVTSSTGDLISI 45
DB 402 EVAKENVQNVKYSNLEIKXHYNDQINIDNTEKEAKQNYQDFKEHMKTIIPPNEWKYQK 461

QY 46 PSELENIPSENYFOSAI--KSGFIKV---KKSDEYTFATSDADNHVTWVDQEV---I 97
DB 462 PSIEIKIMKDE---FLSKYKNYNDPKVYKEVSEHNKFTLTWNKIKTEVSDIEIKKYE 518

QY 98 NKASNNKIRLEKGRUYQIKIQORENPTKGLDFKLYWTDSONKKEVSSNLOLPELK 157
DB 519 NKFNDSKSLINETKK--SIEEYQINILTKVDD---YIKVGLNTNELITNCHNQTLK 573

QY 158 QKSNRKK--KRSTAGPTVDRDNDGIDP-----SLEVEGYTVDVKNKTFELSP 205
DB 574 DKLNQNKITKETSNDKIDYDFENILDTKKTELETKTGSLNNHSENNKELLTYFD 633

QY 206 WISNIEHKKGLTKYKSPKASTADPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHD 265
DB 634 LKANLGNKENMLYKOFNEK-----ERAVEDIKKNVDINKIVSN-----IEIT 677

QY 266 MENILSKNEDQSTQNTDSTRITISKNTSTSRHTSEVHGNAEVHASFDDIGGSVAGFS 325
DB 678 IYTSIYNIND-----TENB---IGKSIELLNTKYLE-----KVKANVT 713

QY 326 NSNSTVAI--DHSLSLAGERTWAGTGLTADTARLNANIRYNTGTAPIYVLPVTS 384
DB 714 NLNEIKELKXDYDFQPGK-----EKNIKYPDEN--KIKNDIDITLQK 754

QY 385 LGKN--QTLATIK--KENQLSQI-----LAPN-----NVYP-----SKNLAPIALNAQK-- 425
DB 755 IDKSIEFTLTIKKNSNHNDEIKGQIDKLVKPNKTFMNEDEPKIEKKIENIVEKIDKKK 814

QY 426 -----DFSSTPIIMYNOFLEKTKQLRLDQVYGNIAIYNFENGVRVVDVTSNWS 479
DB 815 NIYKEIDKLLNEISKIENDKTSLEKLNINLSYKSLGNLFLQOIDEKKKAETHIKAME 874

QY 480 V-----LPQIQTETARIIFNGKDLNL-----VERIAAVNPSDP-----LETT-----K 518
DB 975 AYIDDLNIIKKKQEI---EKEMNINMDIKQDILHEKMKALNTSHDDYKITYHTTSKNHEK 931

QY 519 PMVTLKEALKIAFGFNEP---NGNLQYQGKDITEFDNFQDQTSQNIKNQLAELNATNY 575
DB 932 ISDTRKNSLKIQDFSESYINDIKKELEKNVLE-----SQNNNTDINQVLSKIE--NIY 984

QY 576 TVLDKIKLNKAKNILLRDKFP--HYDRNNIYAGADESVKKEAHREVINSSTEGLLNIDK 633
DB 985 NIL---KLNKIKKIIDKVEYTDIEKON-----KKINAELSNS-----1020

QY 634 DIRKILSGYVEIETDEGLKE-----VINDRY---DMLNISLRQDGKTFIDFKKYN 683
DB 1021 --EKI-----ITQLKENSLEKCSQKSIKSTIDNNVSECIKNITNL-----KTIYVNEKN-- 1069

QY 684 KLPAYISN--PNYKVN-----YAVTKNTIINPSENGDSTN--GIKKILIFSKK 730
DB 1070 -INTYFKNABEYNQNSLNFNNIEMADTKSYILNKKNGNTNNTDYNIKELKEHKK 1126

RESULT 7

S05603
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (s
N;Alternate names: gp195 surface antigen
C;Species: Plasmodium falciparum
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jun-2000
A;Reference number: A45521; MUID:91101660; PMID:2270106
A;Accession: S05603; S04850
R;Myler, P.J.
submitted to the EMBL Data Library, April 1989
A;Reference number: S05603
A;Accession: S05603
A;Molecule type: mRNA
A;Residues: 1-1639 <MYL>
A;Cross-references: EMBL:X15063; NID:g9896; PIDN:CAA33163.1; PID:g9897
R;Myler, P.J.
Nucleic Acids Res. 17, 5401, 1989
A;Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plas
A;Reference number: S04850; MUID:99345116; PMID:2668887
A;Accession: S04850
A;Molecule type: mRNA
A;Residues: 1504-1639 <MYL2>
A;Cross-references: EMBL:X15063
C;Superfamily: major merozoite surface antigen
C;Keywords: glycoprotein; merozoite; surface antigen
F;1-19/DNA: signal sequence #status predicted <SIG>
F;20-1639/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 4.9%; Score 185.5; DB 2; Length 1639;
Best Local Similarity 20.5%; Pred. No. 0.21;
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

QY 5 ENRL-LNESESSQGLLYYFSDINFCAPMVVTSITGDLSPSELENIPSENYFQSA 63
DB 927 ENILSLGKNKNTYQELIGKQKSSB-NF-----YEKILKDSDFYNE 965

QY 64 IWSGFIKVKSDYTFATSDADNHVTWVDQEVINKASN-----NKIRLEKGRUY- 114
DB 966 SFTNFVSKADD-----INSLNDESKRKKLEEDINKLTKTQLSFDLYNKKYKLERLFD 1020

QY 115 -----QIKYQORENPTKGLDFKLYWTDSONKKEVSSNLOLPELKOKSSNSKKRS 168
DB 1021 KKTVGVYKMQIKKLTLLKEQLESKL--NSLNNPKHVL--QNFVSFVNKKGEAIEAETEN 1076

QY 169 TSAGPTVDRDNDGIDPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKY---KSSP-- 223
DB 1077 T-----LENTKILLKHY-----KGLVKYNGESSPLK 1103

QY 224 ---EKWSTADPYSDFE-----KVTGRIDKNVSPPEAR-----HPLVAAYPIVHD 265
DB 1104 TLSEESIQTEDNYASLENFKVLSKLEGLKDNINLEKKLSYLSGLHLHIA-----E 1156

QY 266 MENILSKNEDQSTQNTDSTRITISKNTSTSRHTSEVHGNAEVHASFDDIGGSVAGFS 325
DB 1157 LKEVKKKN---YTGNSPSENNT-----DVNNALSKYKFLPEGTDVAIVVS 1200

QY 326 NSNSTVAI DHSLSLAGERTWAGTGLTADTARLNANIRYNTGTAPIYVLPVTT---- 381
DB 1201 ESGSDTLTLEQSQPKKPPASTHVGAES--NTITTSQ--NVDDDEVDDVIIIVIFGESEEDYDDL 1256

QY 382 -SLVLGKNQTLATIKAKENQLSQILAPNYYPSKNLAPIA---LNAOKDFSSPTITWNYN 437
DB 1257 GQVTVGAIVPSVI---DNILSKI---ENEYEVLYLKLPLAGVYRSLSKQLENNVMTFNVN 1310

QY 438 -----QFLEKTKQLRLDQV--YCNIAI-----YNFENGVRVVDVTSNWSSEVL 481
DB 1311 VKDILNRSFKNFKNV--LESDLIPYKDLTSSNYVVKDPYKFLNKEKDKFLSSYNIYK 1369

QY 482 PQIQTETARIIFNGKDLNLVERRIAANVPSDPL-----ETTKPDM--TLKEALKIAGF 533
DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSKSDLSIKKYLINDQGE 1410

QY 534 NEPN-----GNLQYQKDI TE-----PDFNFDQDQTSQNIKNQLAELNATNY 576
DB 1411 NEKVLPLANNIETLYKTWNKDILDFVHLEAKVLNITYEK---SNVEVKIKELN--YLKT 1465


```

858 ANTYNIL-----KLNKIKTIIDKKVEYTSIEKKNKQINDELNNSEK 899
426 -----DFSSTPITVNYQFLE-----LEKTKQLRLDITD-----QVYGN 458
900 VIKKIEGDLSLKECRSKINSTLDDKIDECIKNINVLKKNILNEETNITWHFKNAEYKN 959
459 IATYNFENGRRVDTGSNWSEVLPIQIETTARIIFNGKOLNLVERIAAIVNPSPDLETTK 518
960 IVLSNFNN-----IEMADNKSQYILEIKONGT---NDHDYNIKELK-SHKDKSNGYK-TE 1010
519 PDMTLKALKIAFGFNEPNLNQYQGD-----ITEFDNFDDQTSQNIKNQLAE 568
1011 ADQNKRAIQ-----NKELEFYKEEVTVLLNKYYAVELKNKFD-KTKDNKSQIIKE 1061
569 LNAITNTYVLDTKIKLNAKQY-----ILRDKRFHYDRNNIAGVDESUVKEAHREVIN- 621
1062 IKDAHNYCTLESKSGSKQWMEIKNEKIHIEDEVANDKSNKAITSIKVSVEPKTKIIKI 1121
622 ---SSTEGLLNIDKDIRKILSGVIVEIEDT-----EGLKEVINDRYDMLNTSSLR 669
1122 NEIRTSDDCKETNLEKQISNLSIDTQTKLTENGKQLKLEEL-----LESLK 1172
670 QDGKTFIDFKYNDKLPLYSN-----PNKYNVYAVTKENTIINPSENGDTSTNG 720
1173 KQKNKNIEDQKDELDEVNSKIKNIENTVQHKKNYEIGI--VEKINSIAKTNKQIESTKE 1230
721 IKK 723
1231 LIK 1233

RESULT 9
B64635
toxin-like outer membrane protein HP0922 - Helicobacter pylori (strain 26695)
C/Species: Helicobacter pylori
C/Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
R/Accession: B64635
R/Tomb, J.P.; White, O.; Karlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori. A/Reference number: A64520; MUID:9739467; PMID:9252185
A/Accession: B64635
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-2529 -TOM-
A/Cross-references: GB:AE000602; GB:AE000511; NID:g2314060; PIDN:AAD07969.1; PID:g23140

Query Match 4.9%; Score 184.5; DB 2; Length 2529;
Best Local Similarity 21.0%; Pred. No. 0.44;
Matches 158; Conservative 94; Mismatches 261; Indels 241; Gaps 36;

QY 123 ENPTEKGLDFKLYWTDSONK---KEVISSDNLQLPELKOKSSNRKKSSTSGAGTVPDR 178
DB 729 QNPAPESVSGVYWL--QNTYKNKGIYID----PNLSQGSQSGNTLSTYITANLF--- 779
QY 179 DNDGIPDPSLEVEGYVDVKNRTFLSP---WI-----SNIHEKKGLTKYK 220
DB 780 ---GRSPVNIQNTGLIIGNNTSVNSGLIIGHGGFGVITGTFSAANIY---LNNF 832
QY 221 SSPEKWT-----ASD-----PYSDFEKTGTGRIDKNVSEARHPVAAPIV 262
DB 833 KTEGVNSDGGGANITFKASDNITMDGLNYNDAETVTKMIQTGAS---QHSYATFDALN 889
QY 263 HVDMENIILSKNEDQSTQNTDSETRTTSKNTSTSRHTSEVHGNAEVHAGFPFDIGGSVA 322
DB 890 NISVNTSSFS-----DMTWGFFSFAKNISF-----NASF-----S 921
QY 323 GFSNNSSTVA---IDHSLSLAGERTWAETWGLNTADTARINAIRVYVTGTAPIYVL- 378
DB 922 GFTNPGGSSVISANATNSLFSINR-----LNGGAVYNLQANSLFN-NTQAFVNLV 973

```

379 QY -----PTTSLVLGKQNTLATIKAKENQSLAPNNYPSKNLAPIAL--NAOKDFS 428
974 DB SGTENFNATLQGNFTLSS-----QSLNFGDDTTLQNNANITLGNKSOAFK 1025
429 QY STPIITMYNQFLELEKTKQLRLDTPQVYGNIAIYVNFENGRVYVDTGNSWSEVLPOIETT 488
1026 DB NS-LTLDNNSNLSLDNQSVLANNTSAFVNOQASLIYNGS-----QATF 1068
489 QY ARIIFNGKDLNL-VERRTAANPSPLETTKPDWTLKALKIATGFNPNLQYQKDI 547
1069 DB NSLFFNGGTLNASSKLNASAFSNNTT---INLDDSVLSASNTSSLNANINFGASQ 1125
548 QY TEF-----DFNFDQOTSQNIKNOLA-----E 568
1126 DB ADFGNGTIIDTASFNDSASSLNFNNLTANGALNFNGVTPSLTKALMSVSGQFVLGNGD 1185
569 QY LNATIIYVULDKLNKAKNNIIRKRP-----HYDRNNIAV 605
1186 DB INLSDI-NIFDNITTSVTYNIILNAQKGITGSGANGYEKILFYGMKIQIATYSNNNIQT 1244
606 QY GA-----DESUVKGAHR-----EVIN--SSTEGLLNIDKDI-----RKILSGYIVE 645
1245 DB WFINPLNSQIIQBSIKNGDLTIEVLNPNFASNTIENIAPELYNQASKQNPFGYSYD 1304
646 QY IEDTEGLKEVINDRYDMLNSSL-----RQDKTFIDFKYNDKLPY-----ISNPNY 694
1305 DB YSDNOA-----GTYILTSNIKGLFTPKGSQTPAQFTYSPFNQPLSSLNIYKGFSEN 1359
695 QY KYNVYAVTTKENTIN--PSENGDSTNGIKKIL 725
1360 DB KULGLILSONSATIKEMIESNQLDNITN-INEVL 1392
RESULT 10
E90598
membrane nuclease, lipoprotein [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: E90598
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A>Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: E90598
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1125 <KUR>
A:Cross-references: GB:AL445566; PID:gl4090108; PIDN:CAC13866.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU 6930
A:Genetic code: SGC3
Query Match 4.9%; Score 183.5; DB 2; Length 1125;
Best Local Similarity 19.1%; Pred. No. 0.15;
Matches 165; Conservative 162; Mismatches 330; Indels 207; Gaps 40;
QY 3 KQENRLNESESSQGLGYFSDLPFQAPMVYTSSTGDLGIPSSSELENIPSENYFQS 62
DB 120 ENDNSVNNKENSSK-----NDEN-----LITKVLRLGHVNV-LNQSGNITPKNLAIK 168
QY 63 AIWS-----GFIKV-----KGSDEYTFATSDNHNVTMWYDQGE---VI 97
DB 169 VILHNKLDVLGLTEITNNGVKTIIINELNKSDDKNYIVSLKRGKTGTGSSGQEHVGI 228
QY 98 NKASNSNKIRLE-----KGRLYQIKIQYQRENTEKGL----- 130
DB 229 YK---ENKLTLESDDKEKGRPYENKLM---DDPFKQKQKIDFVRPFGVKFSTKGN 282
QY 131 --DFKLXW--TDSQNKKE---VISSDNLQFLFKQKSSNRKRKSTRSAGTVPDRDNDGI 183
DB 283 KNDFTVFMHSDAPGVKKEGELISAKGYSGQGHKEVAEALRTKEVMEYFDSI-----DGV 337

184 QY PDSLEVEGYT-----VDVKNKRTFLSP-W--ISNIHEK---K 214
338 DB NNELFPNGDTNKLGNKAKAFEPLLQSGYKSLIKOVKENATSLAQRWGEYANHDKIFYK 397
215 QY GLTKYKSSP--BKWSTADPYSDPEKVTGRIDK-NVSPKARP-----LVAAYPI 261
398 DB GDLKVENSGFYDLWKVFDNLLNKBEFTKVBEYSSKKAKDKYKGEYSYVILHISDHTI 457
262 QY VHVDMENIILSKNEDOSTQNTDSETRTISKTSTSTRTHTSEVHGNAEVHASFDFIGGSYS 321
458 DB VYTDL--IILQKDD---QNKSENK--DENENDSKQLDKPTNNEQNQNTQDDSKKIS 509
322 QY AGFSNNSSTVAIDHSLSLAGERTWAERTMGLNTADTARLNANIRVYNTGTARIYVNLPT 381
510 DB DASQNNSTTNEKQKLDSDQE---SKNNAIKSQONQDKDNLSSKNDTQSKSESSPOI 566
382 QY SLVLGKQNTLATIKAKENQSLAPNNYPSKNLAPIALNAQKDFSSSTPIITMYNQFILE 441
567 DB NPWLENNQELSHNNGENDDSKQNTSNRQTKN---DLRSEQKNLT----- 611
442 QY LEXTQLRLDTPQVYGNIAIYVNFENGRV--VDTGSNWSEVLPOIETTARIIFNGKDL 498
612 DB ---TKNPSSNSNVETKNETQNNENSSTKKEDITSA-----KTQDSTNENLKNEET 661
499 QY NLYVERI-----AANPSDPLETTKPDWTLKALKI--AFGFNEPENGLOVQGDITE 549
662 DB NQVETKNTNENNSNTNKEENSSTKKEEISKSBSNVNNSSTKQENIYNKKEEISK 721
550 QY FDFNFDQOTSQNIKNQ-LAELNATNIYTVLDKIKLNQKN-----ILIRDKRPHYDRNNI 603
722 DB SESNVNNSNTNTQNCETPETNESQNNVIGKNPNQSLNQNDAIDVSAKKVIGYWNINE 781
604 QY AVGADSVVKEAHPREVNSTEGLL-----LNIDKDIRKILSGVIVEIEDEGLKEVINDR 659
782 DB SVGKSASAKAFVAKVIDHDKLDLVIGIGLVHEETLTKI-----VEENKLSKDSK 834
660 QY YDMLNITSSLPQDQKTF-IDPKYNDKLPYISNPYKVNVAVTKENT---INPSENGD 715
835 DB W--VQVISEKQKQEGFPVNLARYIG--VIVKEN---KFNIESPKQNTNKGHLHYENQPNW 887
716 QY TSTNGIKKI-----LIPSKKG 731
888 DB SSFNTSEKSVYVRPPFGIKFSTKG 911
RESULT 11
T31102
filamentous hemagglutinin 1 - Haemophilus ducreyi
C:Species: Haemophilus ducreyi
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T31102
R:Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.
J. Bacteriol. 180, 6013-6022, 1998
A>Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
A:Reference number: Z20984; MUID:99030326; PMID:9811662
A:Accession: T31102
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4152 <WAR>
A:Cross-references: EMBL:AF057695; NID:G3929017; PID:G3929018; PIDN:AAC79757.1
C:Genetics:
A:Gene: lspA1
Query Match 4.9%; Score 183; DB 2; Length 4152;
Best Local Similarity 22.2%; Pred. No. 1.1;
Matches 187; Conservative 118; Mismatches 321; Indels 218; Gaps 45;
QY 10 NESSESSQGLGYFSDLPNFQ-APMVYTSSTGD-----LSIPSELENIPSENYF 60
DB 162 SRTESTSQ-LVGKLUHAIQKQKAEKILNQVTQDHESNIQGALEAVACKADLIIVN---- 216
QY 61 QSAIWSGPIKVKSGDEYTFATSDNHNVTMWYDQEVINKASN-----SNKIRLEKRLY 114

Db 217 PNGITLGVKNTINTDRFVVST-----DIIPHRENGLLSVNGKVTIDKGGVA 264
 Qy 115 QIKIQY---ORENPEKGLDFKLYWTDSONKEVISSD-----NLQPELKOKSS 161
 Db 265 TNGLSHFVAVARNIDQKG--KITVAKTENQKSVNPANITFAAGSLNVLKTRATPISS 321
 Qy 162 NSRKRSTAGTPVDRNDGIPDSLEVEGTVD-----VKRRTFLSPWISNIHEKGLT 217
 Db 322 GTSR---TSDTPAI--SADSAGSMYGSNFKFVVTDKAGVGHKGIIFSENDINIKWDGNA 377
 Qy 218 KYKSPKSWASTDPSDFEKTGRIDKNVSPERARHPLVAAPVPIVHVDMENIILSKNEDQ 277
 Db 378 SLKEIYAK-----KODILAKDIELTEKGQLOANIKILNSTGKINLNASV 425
 Qy 278 STQWDSSTRIS--KNTSRTSRTHTSEVGNABVHASFFDGGSVSAGFSNSSTVALDH 336
 Db 426 SADNVNVSLENALENASMSANSLDVIVTKIEVRS-----SKVSAGTANIKASNITLDG 480
 Qy 337 SLSLAGERTW-----AETWGLNTADTARLN-----ANIRVYV---T 369
 Db 481 SSVVANKITLVNTNATLNQSKLSAKDMELNVTHENTLNSTKLSAQKANKIKENTLN 540
 Qy 370 GTAPIYNVLPTSLVLGK-----NOTLA-----TIKAKENQ--LSQILAPNN-- 409
 Db 541 GEA-----SLVAEKLIDINAIDKITNNGTITAGLTANITTKALENRDNALILAHQNLN 591
 Qy 410 -----YV-----PSKNLAPIALNAQKDFSS-----TPITMYNQFLELEKTKOL 448
 Db 592 FTVNGSHVYKGDIVSKKAVTFPNSGDFTSNGSKLVDAQNLTNVNANNITQGSSEII 651
 Qy 449 RLDTQVYVNI---ATYAFEN--GRVRVDTGNSWSEVLPOIQTETARIIFNGKDLNVER- 503
 Db 652 -----LHGNTVLNAKGNFTSNGLTMTKELNISIESFI---NAGNLTGKLEHVSNT 702
 Qy 504 -----RIAAVPSPLETKPDMTL--KEALKIAGF---FNEPNGNLOVQCKDITEF 550
 Db 703 TVKNDGLVSIENLNISSKTDFTNNGTLGLEALKIAGGFTNASGSLA--SKNSLDIY 761
 Qy 551 DFNF--DQOTSQNIKQLAELNATNIYTVLDK--IKLNKAKMILIRDKRPHYDRNNIAGVA 607
 Db 762 GNNFTNGTIESVKS---LNTNNVTPINNATIKSYGLVNIISQ--NFTNDSNGTVMSH 816
 Qy 608 D-ESVVKAEHREVINSST-----EGILL-----NIDKDKILSGYIVEIEDTEGLKEVIND 658
 Db 817 DLLNTTSQAN--IIRKNLAGQGLNTAKGNITNDSNSTAIKVLHNSNDIN-----LNA 869
 Qy 659 RYDMLNISSL--RDGKTFIDFKKYNKPLPLVISNPNYKVNVAVTKEINTIINPSENGDTS 717
 Db 870 NKNYVNIIGIYSQAGNISVEAKLHNDVKL-----SGNITTTTK-----SGNATVK 915
 Qy 718 TNGI 721
 Db 916 TNSI 919

RESULT 12

B89921
 hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: B89921
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A:Reference number: A89758; PMID:21311952; PMID:11418146
 A:Accession: B89921
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-6713 <KUR>
 A:Cross-references: GB:BA000018; PID:gl3701232; PIDN:BA842527.1; GSPDB:GN00149

A:Experimental source: strain N315
 C:Genetics:
 A:Gene: ebha

Query Match 4.8%; Score 182; DB 2; Length 6713;
 Best Local Similarity 20.5%; Pred. No. 2.5;
 Matches 171; Conservative 130; Mismatches 299; Indels 234; Gaps 44;

Qy 9 LNESESSOGLLYFSDLNFOAPVWVTS-----STTGDLSPSSELENIPSENQYFQSA 63
 Db 2468 VRQAKSDAKANLG--TLTHLNAQKQDLTSIQEGATTVNGSVNKTQAQDLGAMORLESA 2526
 Qy 64 IWSGFTKVKKSDY-----TFATSADNHVTMVWDQEVINKASNSNKIRLEKRLYQIKI 118
 Db 2527 I--ANKDQTKASENYIDADFTKTAFDNAIT---QAESVYLNKDHGTNKDK-----QAVBQ 2576
 Qy 119 QYORENPTKGL--DFKLYWTDSONKEVISSDNLQPE-----LQKSSNSRKRSTSGAP 173
 Db 2577 AIOQVSTSTENALGNDANLQCAKTEATQAIQDNITQLTNPQKTAQKQVNAQR----- 2628
 Qy 174 TVPDRNDGIPDSLEVEGTVDVKNKRTFLSPWISNIHEK-----KGLTKYKSSPEK 225
 Db 2629 -----VSGVT--DLKNSATSLNNAQDLKQAIGDHTIIVAGNVYTNASDPK 2672
 Qy 226 WSTASDPYSDFEKVTGRIDKNVSPERARHPLVAAPVPIVHVDMENIILSKNEDQSTQNTDSE 285
 Db 2673 QGAYTDAYNAAKNIVNG--SPNVITNAADVTAAATQRV-----NNAETSLNGDTN 2719
 Qy 286 TRII---SKNTSRTSRTHTSEVGNABVHASFFDGGSVSA-----GFSNSNSTVAIDHS 337
 Db 2720 LATAKQAKDALKOMTHLSDAQKQS-----ITGIDTSATQVTVGVSKVKNATNLDNA 2771
 Qy 338 L-----SLAGERTWAEITMGLNTADTARLNATNIYVNTGTAPIYVNLPTSLVLGKNTU- 391
 Db 2772 MNQLRNSIANKDEKVASQPVVDADTDKQNA-----YNTAVTSAENIINATS-----OPTLD 2822
 Qy 392 -ATIKAKENQ-----SQILAPENYYPKSNLAPIA--LNAOKDFSTPITMYN--Q 438
 Db 2823 PSATVQAANQVNTNKTALNGAQLNAKQKETTANIRLSHLNNAQKQDLNTQVTHAPNIS 2882
 Qy 439 FLELEKTKOLR-----DTPQV-----YGNAT-----YNPE 465
 Db 2883 TVNQVTKAEQLDQAMERLINGQDKQVQKSVNFTDADPEKQATYNNAVTAENIINQA 2942
 Qy 466 NGRVRVDTGNSWSEVLPOIQTETARIIFNGKDLNVERRI--AAVNPSPDPLET----- 516
 Db 2943 NG-----TNANQSVBAALSTVTTTQKALNG-----DRKVTDAKNANQTLSTLDNLNN 2991
 Qy 517 -----TKPDMTLKEALKIAGFNEPNGNLQ--YQKDKITEFDFNFD----- 555
 Db 2992 AQKGAVTGNINQAHTVAEVT--QAIQTAQELNATMGNLKNLNDKDTLLGSGNQFADADPE 3049
 Qy 556 -----QOTSQNIKQLAELNATNIYTVLDKIKLNKAKMILIRDKRPHYDRNNIAGVADES 610
 Db 3050 KKNAYNEAVRNAENILNKSTGTNV-----PKDQVEAAMN-----QVNTTKAALNGTON 3097
 Qy 611 VVKEAAREVINSSTGL--LLNIDKDKIRKILSGYIVEIEDTEGLKEVINDRYDMLNISS- 667
 Db 3098 L--EKAKQHANTAIIDGLSHLTHNAQKALQQLVQOSTTVAEAQNEQKAN-----NVDA 3149
 Qy 668 ---LRQ---DGKTFIDFKKYNKPLPLVISNPNYKVNVAVTKEINTIINPSEN 713
 Db 3150 MDKLRQSIADNATKQNKQNTD-----ASPN--KDAYNNAVTTAQGIIDQTTN 3196

RESULT 13

S68218
 botulinum neurotoxin type A, nontoxic/nonhemagglutinating - Clostridium botulinum (stra.
 C:Species: Clostridium botulinum
 A:Variety: strain NIH
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 02-Jun-2003
 C:Accession: S68218; S74301
 R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.

FBBS Lett. 376, 41-44, 1995
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin component
A:Reference number: S67988; MUID:96096783; PMID:8521962
A:Accession: S68218
A:Molecule type: DNA
A:Residues: 1-1193 <FUJ>
A:Cross-references: EMBL:D67030; NID:G2160224; PIDN:BAAL1050.1; PID:d1011710; PID:g11324
A:Experimental source: strain NIH
A:Accession: S74301
A:Molecule type: protein
A:Residues: 1-13;145-155 <FUI>
A:Experimental source: strain NIH
C:Genetics:
A:Gene: ant
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin
F:1-1193/Product: botulinum neurotoxin type A, nontoxic/nonhemagglutinating #status expected

Query Match 4.8%; Score 180; DB 2; Length 1193;
Best Local Similarity 19.6%; Pred. No. 0.25;
Matches 191; Conservative 131; Mismatches 319; Indels 294; Gaps 46;

QY 42 DLISIP--SSELENTPEQNQYFQSAIWSGFIKVK--KSDYTFATS-----ADNHVT 88
DB 241 DLIVPYRLRSELENTPEQNQYFQSAIWSGFIKVK--KSDYTFATS-----ADNHVT 88

QY 89 MWVDDDEVINKASNSNKIRLEK-----GRLYQIKIQY-----QRENPTKGLDF 132
DB 301 IYETEIEGNAIGNDIKLRKQKFRININDIWEALNLYFSKFFSMMWPRFNALKHFR 360

QY 133 KLYW-----TDQNKKEVISS-----148
DB 361 QYKIKIDYPENYSINGFVNGQINQAULSDRNDQIINKEEIIINLLNGNVSLSMESNYG 420

QY 149 -----DNLOLP-----ELKOKSNSRKRSTSGPTVPDRDNDGIPDSLEVRGY 192
DB 421 DGLKSTVDPEFSYKIKIPYRAREYEHNNDSGLNNVIGVI-----DNIFEIDVNPY 474

QY 193 TVD-----VKNKRTFLS--PWISNHEKGLTKYKSPKSTASDPVSDFKVTG 241
DB 475 KENCDFSPVQKITSTREINTNIPWPIY-----LOAQNTNNEKFSLS-----SDFEVVS 525

QY 242 RIDKNSVSPARHPVAAPIYVHVDMENII--LSKNDQSTQNTDSB---TRTISKNTST 295
DB 526 SKDKS-----LVYFSLNVMFYLDISKNSPIDTDKKYILWLEIFRNTYSF 571

QY 296 SRTHTSVHGNAEVH-----ASFFDIDGGSVSAFNSNS--STVAIDHLSLAGE 345
DB 572 DITATQENTNCGINKVVTWFGKALNLTSDSFVEEFONLGAISLKNKNSLWPIES 631

QY 346 WA---ETWGLNTADTARLNANIRYVNTG--TAPIYVLPY-----TSLVLGKNT 390
DB 632 YEIPNDMLGFLDLNLEKLENIYSKNTAYPKKIYYNPLDOWTQYYSQYFDLICHAKRSV 691

QY 391 LA---TIKAKENQLSQILAPNNYPSKXLAIAL---NAQKPFSS--TPITMN-----435
DB 692 LAQETLIKRIIOKKLSYLGINSISSDNALMLNTTLNLRDISNESQILAMNVDNFLN 751

QY 436 -----YNQPLEKTKQLRLDQVYG--NATYFNGRVRVDTGNSWSEVLQ 483
DB 752 AALCVFESNYPKFISF-----MEQCINNINIKTEF-----IQKTNINEDEKLQ 798

QY 484 IQETTARIIFNGKDLNLVERRIAAVNPSPLETKPDMTLKEAL-----KIAFGNPNGN 539
DB 799 INQN-----VFNSLDFFELN-----IQNKSLFSSSETALLKEETWPELVLYAFKEPGNN 849

QY 540 -----LQVQGDKI--TEFPNFD-----QQTSONIKQLAEALNATYTV-----577
DB 850 VIGDAGKNTSIEY--SKDIGLVGINSIDALYLANGSNQISFSDFPENGLTNSFSYFWL 908

QY 578 --LQKIKLNACKY-----NILIRD---KRFHY---598
DB 909 RNLGKDIKSLKSKEDNCGEIYFQCTGLVFNMDNSNGEKNIYLSVDSNNSWHYITI 968

QY 599 --DR--NNIATGADSESVV--KEAHREVINSTBGLLNIDKDIRKILSGYVIEIEDEGLK 653
DB 969 SVDRLEKEQLLIFIDNLANVESIKILNLYSNIISLSENNPSYIEGLTILNKPTTS-Q 1027

QY 654 EVINDRYDMLNLISSLRQDQKTFIDPKKYNKDLPLVISNPKYN---VYAVTKENTI---707
DB 1028 EVLSNYFEVLNNSYIRDSNEERLE---YNTKYQLY---NYVFSKPICEVKQNNIYLT 1080

QY 708 INPSENGDTSTNGIKILIFSKGY 732
DB 1081 INNTNNLNLOASKFKLLSNPNKQY 1105

RESULT 14
C97033
Uncharacterized protein, probably surface-located [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: C97033
R:Nolling, J.; Brenner, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C.
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97033
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1227 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79054.1; PID:g15023995; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1080

Query Match 4.7%; Score 179; DB 2; Length 1227;
Best Local Similarity 18.9%; Pred. No. 0.29;
Matches 166; Conservative 136; Mismatches 288; Indels 288; Gaps 44;

QY 37 SSTTGLSISPSSELENTPEQNQYFQSAIWM--SGFIKVKKSDYTFATSADNHVTMWVDDQF 95
DB 14 SVTTSVALLSKPAKAYAADNSVKYSISNSINVKGKE-----VQKE 58

QY 96 V--INKASNSKIRLEKRLYQIKIQVORENPTK-----GLDFKLYWTDQNKKEVI 146
DB 59 VVANKNDNSN--KVSSSENQVS---NKENSNEKVSSEISQINKVNLQVQNNKSVL 114

QY 147 SSDNLQLPKQKSSNRKK-----RSTSGPTVPDRDNDGIPDSLEVEGYTVDVKNKT 201
DB 115 AASNVDEVKINGNVTSDYTAIGETKVKPDTLDIINKAIVDARSAG--TDLSEVEI 172

QY 202 F-----LSPWISNI-----HEKKG-LTKYK 220
DB 173 YDIVSQTAETAEAFRINDGVANVSDYTLGATFVNDANLDSVNKYFYHKRYATVTKFX 232

QY 221 SSPKWNSTA-----SDPYSDFEKVTGRIDKQVSPARHPVAA 258
DB 233 DVATKTSNALKNHNGOGGETDYTALEVSGVTGYLDL-----VKNKIYKEQ-----280

QY 259 YPIVHVDMENIILSKNEDQS--TQNTDSETRTISK--NTSTSTRTHTSEVHGNAEVHASF 316
DB 281 -----NKGRLTITISDSAASTIARINTALDN-----MDA 311

QY 317 GGSVSAGSFSNSSTVAIDHSL---SLAGERTWAEFM---GLNTADTARLNANI-----364
DB 312 GVATLEDYQAIGANNVPLQHVADYNSLAMPQORWGVSEADIGINTIYTNINNSGVGTE 371

QY 365 -RYNTGTGAPI-----YVNLPTTSL--VLGKNTLATIKAKENQLSQILAPNNY---411
DB 372 DDYINSHAVDSNEGNIDYDILNANIIEKTKAGQDL--TIPEVANVKEVKTLDFYNHAA 430

QY 412 -----PSKNLAPALNAQKFSSTPIYNNYNOFLEKTKQLRLDQV-----455
DB 431 AGQTTLDQYKNVDENAOVQDDVAT-----LSMDLXTRDCKTKLKALQDKIDSILSLKN 484

```
QY 456 ---YGNATYNFENGRVYRVDGNSWSEVLPOIQTETAEI-IPNGKOLNVERIAAVNP 510
Db 485 INSGIGNIDDIS-----KIQTEAVDASKLBAVNDIDIKKADKORDLTIQEIIRDSVKKT 538
QY 511 SDPLETT-----KPDMTLKEALKIAF-GFNEPENGLOYQKDIITFDNFDOQTQNTK-- 563
Db 539 IDYINSTGNVKGDSVSDYITIGIDGVTEI--NIEFVNERIKESGITI--TIENIKV 593
QY 564 ---NOLAEIN--ATNIYTVLD-----KIKLNAXNNILIRDKRHYD 599
Db 594 IEPVOLSEVYRIVTGVTVVYKTLGINNVNDNNIYINAEKKNKDVKIQIDQTRVD 653
QY 600 R--NNI-----AVGADES-----VKEAHREVINSSTEGLLNIDIRKILSGYIVE 645
Db 654 NTINNIDVINKIGADVLSDYFNGITDVQDILDVYNADLKIQNYKQVDDIIRVEAK 713
QY 646 IEDTEGLK-----EVINRDYMLNLSLRQDGKTFIDFKKNDKPLY-----ISNPYK- 695
Db 714 ISSYBALMRINIGEAUTDDFKALGJTDI-----NOGLLYATTDLQNKYKT 760
QY 696 -----VNVY-AVTKENTIINPSENGDTSTNGI 721
Db 761 ADEVIARVQAQIEIYRALMQIN--LGKATTADYNTLGI 796

RESULT 15
A86827
hypothetical protein yqfG [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: A86827
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jallion, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: A86827
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1072 <STO>
A:Cross-references: GB:AE005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yqfG

Query Match 4.7%; Score 178.5; DB 2; Length 1072;
Best Local Similarity 19.4%; Pred. No. 0.25;
Matches 160; Conservative 142; Mismatches 342; Indels 179; Gaps 30;

QY 2 VKQENRLNSESBSQGLGYFFSOLNFOAPMVVTSSTTGDLSIPSELEN----- 52
Db 300 VEQVDVASSESTQDANSASLYPISEASSVTDNLNSISLSDSSISSQTSQSGASSTA 359
QY 53 ---IPSENQYFQSA---IWSGFIKVKSDYEYFATSAOHVMTVMVDDQEVINKASNKKI 106
Db 360 EISYDSENENSJSSNQINSNSKSEKDSQSLGSSMSSEHNSNSNTNETNNSSEI 419
QY 107 R-----LEKRLYQIKIQYORENPTKGLDFKL-----YWTDSQNKKEVISSDNLQLPK 157
Db 420 TNLPLPSNPTENSNSVDQTSSEASTNSNSISLSPNSISSTSDSESATNSDDFSNVAEVA 479
QY 158 QKSSNRKRKRKTSAGTVPDRNDGI-----PDSLEVEGYTVDVKNKRTFLSPWISNIHEK 213
Db 480 NNSLASVNNSSSVLSSTSTADNLGINQSGSDNLTKD--SSEISTSGAFLS---SNQTS 534
QY 214 KGLTKYKS---SPEKWS-----TASDPYSDFEKVTGRIDKNVSPRARHPLVAAYPI 261
Db 535 EASTNSNSISLSPNSISSTSVLESTSSNFSNVAEVANNGLASVNNSSSVLSSTSTA 594
QY 262 VHVDM-----ENI-----ILSKNEQSQNTQDSETRTISXNTSTSRHTSEVH 304
Db 595 DNLEINQFGSDNLTKDSSBISTSGAFLSSNQTSSEASSNSMSSINSPLSLSLTNSESA 654
```

```
QY 305 GNAEVHAFDIDGGSVSAGFSNSNSTVAIDHSLSL--AGERTWAETMGLTADTARLNA 362
Db 655 TNO-----SNSEATKVNNST-----HSSNILNSGSDSDSDSDSDSDSNL-- 699
QY 363 NIRVNTGTAPIYNVLPPTSLVLGKNQTLATKAKENQLSQ-----ILAPNNTYPSKNLAP 418
Db 700 -----SSPNLETNQITISSKPEVNNISENPVKVSSSNSVQENSTDEH 742
QY 419 IALNAOKDFSSPTITWNYNQFLEKTKQLRLDTQOVYGNIAATYN---PENGVRVYDTGS 475
Db 743 MSTNPKSSISSPISITSSSQOKESQSN---LLNTTEGINNPITFNNSSSENSAASILT-- 797
QY 476 NMSEVLPOIQTETARIIF-----NGKOL-----NLVERIAAVNPSPDPLETTKPDMT 522
Db 798 SYSNNSESSETCGLYISNEAQRDNGSEISHSLPSSNSNNVSSIQSOAILLESSKSTN 857
QY 523 LK-EALKIAGFNEP-NGNLOYQGGDIITEFDNFDOQTSONIKNQLAEINATNIYTVLDK 580
Db 858 KRSSLSIINSTSHPNQEDNQNSSD-----EVKSNNVVESILGQUNLSINKTHMS 909
QY 581 I---KLNAXNNILIRDKRHYDRNNIAGVADSIVKAEHREVINSSTEGLLNIDIRK 637
Db 910 LTSQKLSVITVLPSSKSKVTNEKNENSTVSEKLIKTPQKN-DESONLQITALDLSFNK 968
QY 638 ILSGYIVEIETDEGLKEVIN-----DRYDMLATISLRQD-GKTFIDFKKNDKPLY 688
Db 969 -----EVETMEDSKTVDPKVLNENGRSQNNKTSTIAKDNKVKFFKRSEFSNKL-- 1019
QY 689 ISNPYKVVYAVTKENTIINPSENGDTSTNGIKKILIPSKKG 731
Db 1020 -----DSDNNILKKTIVLLKKG 1036

Search completed: May 3, 2004, 19:41:13
Job time : 17.0855 secs
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 8.2539 Seconds

(without alignments)
4636.784 Million cell updates/sec

Title: US-09-848-909A-8

Perfect score: 3773

Sequence: 1 EVKQENRLNSESSESSQGLL.....TSTNGIKKILIFSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3764	99.8	764	1	PAG BACAN
2	235.5	6.2	204	1	YPBI_BACAN
3	185.5	4.9	1630	1	MSPI_PLAFK
4	185.5	4.9	1639	1	MSPI_PLAFW
5	175	4.6	1858	1	YM67_YEAST
6	169.5	4.5	1230	1	SMC3_YEAST
7	168	4.5	1957	1	SPOF_SCHPO
8	166.5	4.4	2334	1	WAPA_BACSU
9	164	4.3	1803	1	YJL3_YEAST
10	162.5	4.3	979	1	P115_MYCHR
11	161	4.3	1276	1	BXD_CLOBO
12	159.5	4.2	2869	1	RBF1_PLAVB
13	159	4.2	1000	1	S155_YEAST
14	159	4.2	1385	1	FAT1_SCHPO
15	158.5	4.2	1167	1	CAGA_HELPJ
16	157	4.2	1790	1	USO1_YEAST
17	156.5	4.1	1208	1	PCP1_SCHPO
18	156	4.1	1928	1	MY81_YEAST
19	156	4.1	2867	1	RBP2_PLAVB
20	155.5	4.1	2116	1	MY82_DICDI
21	154.5	4.1	1024	1	RIP3_MOUSE
22	154.5	4.1	1487	1	MDS3_YEAST
23	154	4.1	1037	1	KCC4_YEAST
24	153.5	4.1	678	1	YNC7_YEAST
25	153.5	4.1	1233	1	YF16_YEAST
26	153	4.1	1288	1	VACB_HELPJ
27	152.5	4.1	1116	1	YK5A_AQUAE
28	152	4.0	1029	1	RIP3_RAT
29	152	4.0	1420	1	SRB9_YEAST
30	151	4.0	1882	1	Y468_MYCPN
31	149	3.9	918	1	YJMB_CABEL
32	149	3.9	1577	1	HLXA_PROMI
33	148.5	3.9	1271	1	Y338_MYCGE

34 148 3.9 1250 1 BXE_CLOBO Q00496 clostridium
35 147.5 3.9 1225 1 Y305_MYCGE P47551 mycoplasma
36 147.5 3.9 1450 1 N159_YEAST P40477 saccharomyc
37 147.5 3.9 1726 1 MSPI_PLAFK P50495 plasmodium
38 147.5 3.9 2376 1 TAO3_YEAST P40468 saccharomyc
39 146.5 3.9 730 1 GLN3_YEAST P18494 saccharomyc
40 146.5 3.9 1091 1 CIC2_RAT P54290 rattus norv
41 146.5 3.9 1162 1 BXEN_CLOBU Q06366 clostridium
42 146.5 3.9 1388 1 SIR4_YEAST P11978 saccharomyc
43 146.5 3.9 1397 1 CID_DROME P19338 drosophila
44 146.5 3.9 2339 1 RPC1_PLAFK P27625 plasmodium
45 146 3.9 719 1 YM41_YEAST Q03213 saccharomyc

ALIGNMENTS

RESULT 1
PAG_BACAN
ID PAG_BACAN STANDARD; PRT; 764 AA.
AC P13423; Q9FSR7; Q9KH69; Q9RQU2;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins
GN PAGA OR PAG OR PXO1-110. [Contains: PA-20 (PA20); PA-63 (PA63)].
OS Bacillus anthracis.
OG Plasmid pXO1.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89172073; PubMed=3148491;
RA Welkos S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppla S.H.,
RA Schmidt J.J.;
RT "Sequence and analysis of the DNA encoding protective antigen of
RT Bacillus anthracis.";
RL Gene 69:287-300(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=28, 33, BA1024, and BA1035;
RX MEDLINE=99214082; PubMed=10197996;
RA Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;
RT "Genetic diversity in the protective antigen gene of Bacillus
RT anthracis.";
RL J. Bacteriol. 181:2358-2362(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=V770-NP1-R / ATCC 14185;
RX MEDLINE=20359347; PubMed=1089854;
RA Cohen S., Mendelson I., Altbaum Z., Kobiler D., Elhanany E., Bino T.,
RA Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M.,
RA Kronman C., Velan B., Shaffer A.;
RT "Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus
RT anthracis spore vaccines protect against anthrax.";
RL Infect. Immun. 68:4549-4558(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Scerne;
RX MEDLINE=99454943; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
RA Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
RA Martinez Y., Rieke D., Svensson R., Jackson P.J.;
RT "Sequence and organization of pXO1, the large Bacillus anthracis
RT plasmid harboring the Anthrax toxin genes.";
RL J. Bacteriol. 181:6509-6515(1999).
RN [5]
RP DOMAINS.
RX MEDLINE=91332080; PubMed=1651334;
RA Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;
RT "The carboxyl-terminal end of protective antigen is required for
RT receptor binding and anthrax toxin activity.";

RL J. Biol. Chem. 266:15493-15497(1991).
 RN [6]
 RC CHARACTERIZATION.
 RX STRAIN=Stjerne;
 RA MEDLINE=94327640; PubMed=8051159;
 RL Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;
 RT "Anthrax protective antigen forms oligomers during intoxication of
 mammalian cells.";
 RN J. Biol. Chem. 269:20607-20612(1994).
 [7]
 RC CHARACTERIZATION.
 RX STRAIN=Stjerne;
 RA MEDLINE=21129592; PubMed=11207581;
 RL Beaugregard K.E., Collier R.J., Swanson J.A.;
 RT "Proteolytic activation of receptor-bound anthrax protective antigen
 on macrophages promotes its internalization.";
 RN Cell. Microbiol. 2:251-258(2000).
 [8]
 RC TOXIN REGULATION.
 RX STRAIN=Wevbridge;
 RA MEDLINE=94131936; PubMed=8300513;
 RL Koehler T.M., Dai Z., Kaufman-Yarbray M.;
 RT "Regulation of the Bacillus anthracis protective antigen gene: CO2 and
 a trans-acting element activate transcription from one of two
 promoters.";
 RN J. Bacteriol. 176:586-595(1994).
 [9]
 RC MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.
 RX STRAIN=Stjerne;
 RA MEDLINE=95050722; PubMed=7961869;
 RL Singh Y., Klimpel K.R., Arora N., Sharma M., Leppla S.H.;
 RT "The chymotrypsin-sensitive site, PFD315, in anthrax toxin protective
 antigen is required for translocation of lethal factor.";
 RN J. Biol. Chem. 269:29039-29046(1994).
 [10]
 RC MUTAGENESIS OF DOMAIN 4 LOOPS.
 RX STRAIN=Stjerne;
 RA MEDLINE=99185012; PubMed=10085028;
 RL Varughese M., Teixeira A.V., Liu S., Leppla S.H.;
 RT "Identification of a receptor-binding region within domain 4 of the
 protective antigen component of anthrax toxin.";
 RN Infect. Immun. 67:1860-1865(1999).
 [11]
 RC MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.
 RX STRAIN=Stjerne;
 RA MEDLINE=21092804; PubMed=11178978;
 RL Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;
 RT "Trp 346 and Leu 352 residues in protective antigen are required for
 the expression of anthrax lethal toxin activity.";
 RN Biochem. Biophys. Res. Commun. 281:186-192(2001).
 [12]
 RC MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.
 RX STRAIN=Stjerne;
 RA MEDLINE=21138996; PubMed=11554763;
 RL Ahuja N., Kumar P., Bhatnagar R.;
 RT "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are
 required for oligomerization of anthrax protective antigen.";
 RN Biochem. Biophys. Res. Commun. 287:542-549(2001).
 [13]
 RC MUTAGENESIS OF PRO-289.
 RX STRAIN=Stjerne;
 RA MEDLINE=21255689; PubMed=11356563;
 RL Khanna H., Chopra A.P., Arora N., Chaudhry A., Singh Y.;
 RT "Role of residues constituting the 2beta1 strand of domain II in the
 biological activity of anthrax protective antigen.";
 RN FEBS Microbiol. Lett. 199:127-31(2001).
 [14]
 RC MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.
 RX MEDLINE=21125576; PubMed=11222612;
 RL Mogridge J., Mourez M., Collier R.J.;
 RT "Involvement of domain 3 in oligomerization by the protective antigen
 moiety of anthrax toxin.";
 RN J. Bacteriol. 183:2111-2116(2001).
 [15]

RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.
 RX MEDLINE=21269403; PubMed=11131326;
 RA Sellman B.R., Nassi S., Collier R.J.;
 RT "Point mutations in anthrax protective antigen that block
 translocation.";
 RL J. Biol. Chem. 276:8371-8376(2001).
 [16]
 RN MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;
 RP ILE-239; TRP-255 AND PHE-265.
 RX STRAIN=Stjerne;
 RA MEDLINE=22112896; PubMed=12117959;
 RL Chauhan V., Bhatnagar R.;
 RT "Identification of amino acid residues of anthrax protective antigen
 involved in binding with lethal factor.";
 RN Infect. Immun. 70:4477-4484(2002).
 [17]
 RC X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RX MEDLINE=97192099; PubMed=9039918;
 RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;
 RL "Crystal structure of the anthrax toxin protective antigen.";
 RN Nature 385:833-838(1997).
 [18]
 RP REVIEW.
 RX MEDLINE=21428689; PubMed=11544370;
 RA Mock M., Fouet A.;
 RL "Anthrax.";
 RN Annu. Rev. Microbiol. 55:647-671(2001).
 CC -!- FUNCTION: One of the three proteins composing the anthrax toxin,
 CC the agent which infects many mammalian species and that may cause
 CC death. PA binds to a receptor (ATR) in sensitive eukaryotic
 CC cells, thereby facilitating the translocation of the enzymatic
 CC toxin components, edema factor and lethal factor, across the
 CC target cell membrane. PA associated with LF causes death when
 CC injected, PA associated with EF produces edema. PA induces
 CC immunity to infection with anthrax.
 CC -!- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a
 CC protective antigen (PA), a lethal factor (LF) and an edema factor
 CC (EF). None of these is toxic by itself. PA+LF forms the lethal
 CC toxin (LefTx); PA+EF forms the edema toxin (EdTx). PA-63 forms
 CC heptamers and this oligomerization is required for LF or EF
 CC binding. Once activated, at low pH, the heptamer undergoes
 CC conformational changes and converts from prepore to pore inserted
 CC in the membrane, forming cation-selective channels.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- DOMAIN: The molecule is folded into four functional domains. Each
 CC domain is required for a particular step in the toxicity process.
 CC Domain 1 contains two calcium ions and the proteolytic activation
 CC site. Cleavage of the PA monomer releases the subdomain 1a, which
 CC is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is
 CC part of the remaining 63-kDa fragment (PA63) and contains the
 CC binding sites for LF and EF. Domain 2 is a beta-barrel core
 CC containing a large flexible loop that has been implicated in
 CC membrane insertion and pore formation. There is a chymotrypsin
 CC cleavage site in this loop that is required for toxicity. Domain 3
 CC has a hydrophobic patch thought to be involved in protein-protein
 CC interactions. Domain 4 appears to be a separate domain and shows
 CC limited contact with the other three domains: it would swing out
 CC of the way during membrane insertion. It is required for binding
 CC to the receptor; the small loop is involved in receptor
 CC recognition.
 CC -!- PTM: Proteolytic activation by furin or a furin-like protease
 CC cleaves the protein in two parts, PA-20 and PA-63; the latter is
 CC the mature protein. The cleavage occurs at the cell surface and
 CC probably in the serum of infected animals as well; both native and
 CC cleaved PA are able to bind to the cell receptor. The release of
 CC PA20 from the remaining receptor-bound PA63 exposes the binding
 CC site for EF and LF, and promotes oligomerization and
 CC internalization of the protein.
 CC -!- MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were
 CC performed that showed that the residues present in the small loop
 CC of domain 4, and not the ones in the large loop, are involved in
 CC receptor recognition.
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

"Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level.";
EMBO J. 4:3823-3829 (1985).

REVISIONS, SEQUENCE FROM N.A.
Pan W., Tolle R., Bujard H.;
Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (potential).
PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

EMBL; X03371; CA227070.1; -
InterPro; IPR006209; EGF-like.
Pfam; PF00008; EGF; 1.
Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
Transmembrane; GPI-anchor.
SIGNAL 1 19 POTENTIAL.
CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.
DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.
TRANSMEM 1614 1630 MEMBRANE ANCHOR.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE 1630 AA; ADBD8C3CE0A46322 CRC64;

Query Match 4.9%; Score 185.5; DB 1; Length 1630;
Best Local Similarity 20.5%; Pred. No. 0.12;
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

5 ENEL-LNESSSSQGLGYFSPNFOAPMVVTSSTGLSPSSSELENIPSENOYFOSA 63
918 ENILSGKNKIYQELIGQKSB-NF-----YEKILKDSITFFNE 956
64 IWSGFIKVKSDRYTATSDADNHVTWVDQEVINKASNS-----NKIRLEKGRLY- 114
957 SFTNFVKSKADD-----INSLNDESKRKKLEEDINKLTKTQLSFDLYNKYKLERLFD 1011
115 -----QIKIOYQRENPTKGLDFKLYWTDSONKKEVISDNLQLPKOKSSNRKKHS 168
1012 KKTIVGKYKQIKKLTLLKEQLSKU--NSLNPRHVL--QNPSPFNKKKEAEATEN 1067
169 TSAGPTVDRDNGIDPSLVEGYTVDVKNKRTFLSPWISNTHKKGLTKY---KSSP-- 223
1068 T-----LENTKILLKHY-----KGLVKYNGESSPLK 1094
224 ---EKWSTASDPYSDPF-----KVTRIDKNVSPAR-----HPLVAAPYIVHYD 265
1095 TLSEESIQTEENYASLENFKVLKGLKGLKNDLNLEKKLSYLSGGLHLIA-----E 1147
266 MENILSKNEDQSTQNTDSETRITSKNTS*SRTHTEVHGNAVBHASFDDIGSVSAGFS 325
1148 LKEVIKKNQ---YTGNSPSENNT-----DVNNALESYKFLPEGTDVATVVS 1191

326 NSNSTVAIDHSLSLAGESTWAEITMGLNTADTARLNANIRYVNTGTAPINVLPTT--- 381
1192 EGSQDTLEOSQPKKPASTHVGAE---NTIITSQ-NVDEVDVVIIVIFGSEEDYDDL 1247
382 -SLVLGKQTLATIKAKENQLSOLAPNNYPSKNLAFIA---LNAQKDFSTPTTMVYN 437
1248 GQVVTGEAVTPSVI---DNILSKI---ENEVEVLYLKLPLAGVYRSLKQLENNVMTFVN 1301
438 -----QFLELEKTKQLRLDQV-YGNIAI-----YNFENGRVVRVDTGSNWSEVL 481
1302 VKDILNSRNKRNKFNQV-LESDLIPYKDLTSNYYVVDKPYKFLNKKERDKFLSSVNIK 1360
482 PQIQTETARIIFNGKDLNLVERRIAAVNPSDPL-----ETTKPDM-TLKALKIAPGF 533
1361 DSID-----TDINEFA-----NDVLGYVKILSEKYKSDLSIKKIYINDKQGE 1401
534 NEFN-----GNLQVQGDITE-----DFNFDQDTSQNIKKQLAELNATNIYT 576
1402 NEKYLPLNNIETLYKTVDKIDLFVHLEAKVLYNYTEK---SNVEVKIKELN--YLKT 1456
577 VLDKIKLNARVN--ILIRDKRFHYDRNN-----IAGV-ADESVVKEAHREVINSSTGLL 628
1457 IQDLADFKKNNFVGIADLSTDYHNHLLTKFLSTGMVFENLAKTLVLSNLDGNLQGM 1515
629 LNTDKD--TRKILSGYVIEIDTEGLKEVINDRYDMLNLSLRQDGKTFIDFKYNDKLP 686
1516 LNISQHCQVKK-----QCPQNSGCFRLDE-----RECKCLNLYKQEGDKC- 1557
687 LYISNPNYKVNVAVTAKNTIINPSNG-----DTSTNGIKKI 724
1558 --VENFNPTCN-----EN-----NGCDAKAKTEEDSGSNG-KKI 1590

RESULT 4
MSPI PLAPW STANDARD; PRT; 1639 AA.
AC P04933;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens) (PMMSA) [P195].
GN MSP-1.
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN NCBI_TaxID=5948;
RP SEQUENCE FROM N.A.
RX MEDLINE=86014355; PubMed=2995820;
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V., Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T., Freeman R.R.;
RT "Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum merozoites.";
RL Nature 317:270-273 (1985).
RP REVISIONS.
RA Holder A.A.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
CC SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (potential).
CC PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

DR EMBL; X02919; CAA26676.1; --
DR PIR; A24594; A24594.
DR PIR; S05603; S05603.
DR PDB; 1CEJ; 28-MAY-99.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF0008; EGF; 1.
DR Malaria; Mercoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor; 3D-structure.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1639 MERCOITE SURFACE PROTEIN 1.
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 768 768 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 844 844 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 920 920 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;

Query Match
Best Local Similarity 4.9%; Score 185.5; DB 1; Length 1639;
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

QY 5 ENRL-LNESESSQGLLYFSDLPQAPMVVTSSTTGLSPSELENISENQYFOSA 63
DB 927 ENILSGKKNYQELIGQSS-NF-----YKILKSDFTFNE 965
QY 64 INSGFKVKSDEYFATSADNHVTVWDVQVINKASNS-----NKIRLEKGRLY- 114
DB 966 SFTNFVKSADP-----INSLNDESKRKLEEDINKLKTLLQSLDPLYNKVKLERLFD 1020
QY 115 -----QIKIQORENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQSSNSRKRKS 168
DB 1021 KKTGVCKYKQIKKLTLLKEQESKL--NSLNNPHVL--QNFVFFNKKKEAEIAETEN 1076
QY 169 TSAGPTVPDRDNGIPDSLEVEGYTVVKNKRTFLSPMISNHEKKGATKY--KSSP-- 223
DB 1077 T-----LENTKILLKHY-----KGLVKNYNGESSPLK 1103
QY 224 ---EKWSTASDPYSDFE-----KYTGRIKDNVSPPEAR-----HPLVAAPYIVHD 265
DB 1104 TLSESIQETEDNYASLENFKVLSLEGKQKQWLNLEKKKLSVLSGLHLLIA-----E 1156
QY 266 MENILSKNEDOSTQNTDSETRTISKTSTSTRTSTSEVHGNAEVAHFIDIGSVSAGFS 325
DB 1157 LKEVIKKNK---YTGNSPSENNT-----DVNNALSYKKFLPEGTDTVTVS 1200
QY 326 NNSSTVAIDHLSLAGERTWETWGLNLTADTARLNANIRYVNTGTATPIYVLPFT---- 381
DB 1201 ESGSDTLEOSQPKKPASTHVGAES---NTITTSQ-NVDDEVDVIVIFIGSEEDYDDL 1256
QY 382 -SILVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA---LNAQKDFGSTPTIMNYN 437
DB 1257 GQVVTGEAVTPSVI---DNILSKI---ENEYEVLYKPLAGVYSLKKQLENNVMTFN 1310
QY 438 -----QFLELEKTKQLRLDTQV-YGNIAF-----YFNGRVRVDTGNSWSEVL 481
DB 1311 VKDILNSRENKFNKFN-LESGLPIYKDLTSSNVVVKDPYKFLNKEKRDKFLSSYNIK 1369
QY 482 POIQETTARIIPNGKDLNVERRIAANVPSDPL-----ETTKPDM-TLKEALKIAPGF 533
DB 1370 DSID-----TDINFA-----NDVLGYKKILSEKYSKLSLDSIKKYINDKQGE 1410
QY 534 NBNP-----ONLYQVQKDIPE-----FQNFEDQOTSQNKQLAELNATNIYT 576
DB 1411 NEKLYPFLNIIETLYKTVNDKIDLFVHLEAKVLNVTYK-----SNVEVKIKELN--YLKT 1465

QY 577 VLDKIKLINAKMN--ILIRDKRHFHYDRNN-----Iavg-ADESVVKEAHEVINSSTEGLL 628
DB 1466 IQOKLADPFKNMNFVGIADLSTYHNHLLTFLSTGMVFENLAKTVLSNLDGNLQGM 1524
QY 629 LNTDKD--IRKILSGYVIEIDTEGLKEVINRDMYMLNLISSLRQDQKTTIDFKYNDKLP 686
DB 1525 LNTSQHCQVKK-----QCPQNSGCGFRHLDE-----REECCLLNYKQEGDKC- 1566
QY 687 LVISNPNYKVNVAVTKEINTIIPNSNG-----DTSTNGIKKI 724
DB 1567 --VENFNPTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1599

RESULT 5
Y67 YEAST
ID Y67 YEAST STANDARD; PRT; 1658 AA.
AC Q036J; Q04988;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 187.1 kDa protein in GUA1-ERG8 intergenic region.
GN YMR219W OR YMR261.13 OR YMR959.01.
OC Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OK NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII."
RL Nature 387:90-93 (1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; Z49809; CAA89934.1; --
DR EMBL; Z49939; CAA90190.1; --
DR PIR; S55101; S55101.
DR GerMOnline; 142894; --
DR SGD; S0004832; ESC1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
KW Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;

Query Match
Best Local Similarity 4.6%; Score 175; DB 1; Length 1658;
Matches 160; Conservative 134; Mismatches 285; Indels 272; Gaps 40;

QY 1 EVQQRNLLNES-----SQGLLYFSDLPQAPMVVTSSTTGLSPSELENIPSEN 57
DB 290 EMELEDDIVESDAEKDESGAEGTSHS-VDFSKYQMPRTDNTPKIVIKYSDHKVHQ 348
QY 58 QYFQSAIWS-GFKVKKSDP-----YTFATSDN---HVTMWVDDQEVINKASNSNK 105
DB 349 RYSEDGAFDFGVSNI SVDDDESESAESYSAENAVYHNEHEDLDKELIEDSSDS 408
QY 106 IRLEKRLVQIKIQYQRENPTKGLDFKL-----YWTDSQNKKEV 145
DB 409 -----ESQAQSEQSGEDDFYKMKNEKXSTSEETN*SESQDQGFADKATYKNKV 459
QY 146 ISSDNLQLPEL-----KQSSNSRKRKSTSGAGTVPDRDNGIPDSLEVE 190

460 EQENDEPEKDDIIRSSDKNFHGNKNSSEVSENVLENEDTPAIVERENQ-IND---VE 515
191 GTVDVVKKRTPL---SWISNIHR--KKGLTKYKSS-----PEKWTASD----- 231
516 GYDVTGKSVESDLHSP--DNLYDLAARAMLQFOQSRNSNCPQKEQVSSYLGHSNGS 573
232 -----PYSDEPKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQ 277
574 NLSGRSLDESEQIPLKDF--TGENNNKLTD-RGDLSSS---VEIEVEKV---SEKK 622
278 STQNTDSEPTTSKNTSRTHTSEVHGNAEVHAFDIDGGSVAGFSNNSSTVAIDHS 337
623 LDGSTEKEVLPLSTDTTIN-----NSLGNEDSIYSLDDA 658
338 LSLAGERTWAETMGLNTA-----DTARLNARIVNTGTAPIYANVL 378
659 DAISENLTVPLMEIKTKYEVISEVSYSTSVEDNTVAMPPOVEY-----TSFPMND- 713
379 PTTSLVLGKNQTLATIKAKENOLSQILAPNNYPSKNLAPIALNAQKDFSSPTITM- 434
714 PFNSL-----ND---DYKKHDLKSTLAA-----LAPFTKKDABFVAGVTKSLT 758
435 ----NYNQFLELEKTKOLRLDVOVYGNLATYNFENGVRVDTGNSWSEVLPOIQTAR 490
759 STSGHTNIHTSKETKQVS--DUDESTENVTFNENTG-----DENKQSKNFPFVANSTDK 813
491 IIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAFGNPNGLQVQKDIIEF 550
814 STEDNTD---EKYFSAINTVN---VTGDSCEDIETASVVEE---NLRYCEKDMNEA 862
551 DFNFP-DQNTSON---IKNQLA-----ELNATNIYTVLDKTLNKAQWILLRDRFHYD 599
863 EMSGDECVKQNDGSKTQISFTDSDPNFQESNDNTEFSSTK-----YK 907
600 RNNIAGADESVVKEAHR-EVIN-----SSTEGLLNIDKDIRKI 638
908 VRNSDLEDDSLAKELTKAEVVKLDEESEDSEYQDYADPEPNDGSENMIVKTKK- 966
639 LSGYVIEIETGLKEVINDRYDMLISSLRQDKTFLDFKYNKDLKPLVISPKNVKNV 698
967 -----DTLGIVPEPEKVN-----KVHEETUFEANVSSSVNV 999
699 YAVTKENTIIN 709
1000 QNKDMHTDVIN 1010
RESULT 6
SMC3_YEAST STANDARD; PRT; 1230 AA.
AC P47037;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Structural maintenance of chromosome 3 (DA-box protein SMC3).
GN SMC3 OR YJL074C OR J1049.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=W303;
RX MEDLINE=97474309; PubMed=9335333;
RA Michaelis C., Ciosk R., Nasmyth K.;
RT "Cohesins: chromosomal proteins that prevent premature separation of sister chromatids";
RN Cell 91:35-45 (1997).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=96208490; PubMed=8641269;
RA Galibert F., Alexandraki D., Baur A., Boles E., Chalwatzis N.,

Chuat J.-C., Coster F., Cziepluch C., De Haan M., Domdey H., Durand P., Enlian K.-D., Gatus M., Goffeau A., Grivell L.A., Huangmann A., Herbert C.J., Heumann K., Hilger F., Hollenberg C.P., Huang M.-E., Jacq C., Jauniaux J.-C., Katsoulou C., Kirchthar L., Kleine K., Kordes E., Koetter P., Liebl S., Louis E.J., Manus V., Mewes H.-W., Miosga T., Obermaier B., Perez J., Pohl T.M., Portetelle D., Pujolet A., Fumelle B., Ramezani Rad M., Rasmussen S.W., Rose M., Rossau R., Schaaff-Gerstenschlaeger I., Smits P.H.M., Scarcez T., Soriano N., Tovan D., Tzermia M., Van Broekhoven A., Vandenbol M., Wedler D.H., Von Wettstein D., Wambutt R., Zagulski M., Zollner A., Karpfinger-Hartl L.;
"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome X.";
EMBO J. 15:2031-2049 (1996).
[3]
SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1678;
RX Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[4]
IDENTIFICATION IN A COHESIN COMPLEX WITH SMC1, IRR1 AND MCD1, AND INTERACTION OF THE COHESIN COMPLEX WITH SCC2.
MEDLINE=99145468; PubMed=9990856;
Toth A., Ciosk R., Uhlmann F., Galova M., Schleiffer A., Nasmyth K.;
"Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication.";
Genes Dev. 13:320-333 (1999).
[5]
IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3, MCD1 AND IRR1, AND STRUCTURE.
MEDLINE=21980169; PubMed=11983169;
Haering C.H., Loewe J., Hochwagen A., Nasmyth K.;
"Molecular architecture of SMC proteins and the yeast cohesin complex";
Mol. Cell 9:773-788 (2002).
-!- FUNCTION: Involved in chromosome cohesion during cell cycle and in DNA repair. Central component of cohesin complex. The cohesin complex is required for the cohesion of sister chromatids after DNA replication. The cohesin complex apparently forms a large proteinaceous ring within which sister chromatids can be trapped. At anaphase, the complex is cleaved and dissociates from chromatin, allowing sister chromatids to segregate.
-!- SUBUNIT: Cohesin complexes are composed of the SMC1 and SMC3 heterodimer attached via their hinge domain, MCD1/SCC1 which link them, and IRR1/SCC3, which interacts with MCD1. The cohesin complex also interacts with SCC2, which is required for its association with chromosomes.
-!- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatin. Before prophase it is scattered along chromosome arms. At anaphase, the MCD1 subunit of the cohesin complex is cleaved, leading to the dissociation of the complex from chromosomes, allowing chromosome separation.
-!- DOMAIN: The flexible hinge domain, which separates the large intramolecular coiled coil regions, allows the heterotypic interaction with the corresponding domain of SMC1, forming a V-shaped heterodimer. The two heads of the heterodimer are then connected by different ends of the cleavable MCD1 protein, forming a ring structure (By similarity).
-!- SIMILARITY: Belongs to the SMC family. SMC3 subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Y14278; CAA74655.1; -;
DR EMBL; Z49349; CAA89366.1; -;
DR EMBL; X88851; CAA61313.1; -;
DR PIR; S56850; S56850.
DR GeneOnline; 141688; -;

DR SGD; S0003610; SMC3.
DR GO; GO:0007151; P:sporulation (sensu Saccharomycetes); IMP.
DR GO; GO:0007130; P:synaptonemal complex formation; IMP.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR003405; SMC_C.
DR InterPro; IPR003395; SMC_N.
DR Pfam; PF02483; SMC_C; 1.
DR Pfam; PF02463; SMC_N; 1.
KW Mitosis; Cell cycle; Chromosome partition; ATP-binding; Coiled coil;
KW Nuclear protein.
FT NP BIND 32 39 ATP (POTENTIAL).
FT DOMAIN 172 482 COILED COIL (POTENTIAL).
FT DOMAIN 483 684 FLEXIBLE HINGE.
FT DOMAIN 685 1041 COILED COIL (POTENTIAL).
FT DOMAIN 1126 1161 ALA/ASP-RICH (DA-BOX).
SQ SEQUENCE 1230 AA; 141336 MW; B152D88F7780341F CRC64;

Query Match 4.5%; Score 169.5; DB 1; Length 1230;
Best Local Similarity 19.8%; Pred. No. 0.5; Mismatches 355; Indels 247; Gaps 38;
Matches 182; Conservative 136;

QY 1 EYKQENRLNSESSESGLLGYFFDLNFPQPMVVT--SSTTGDSLSPSSELENIPSNQ 58
Db EMEQERKELEKYNELERNKIYQFTLYDRELNEVINQMERLGDYN-----NTVYSSEQ 262
QY 59 YFQ-----SAIWSGFIKVKSDYEYTFATSDAHNHYTMWVDDOEVINK--- 99
Db YIQELDKREDMDIQVSKLSSI-EASLKIKNATDLOQAKLRESEISOKLTNNVKKIDVQ 321
QY 100 ---ASNSNIRLEKGLYQIK-IQYQRENPTKGLDFKLYWDSONKKEVISEDNLQLP 155
Db 322 QQIESNEEQRLNDSATLKEIKSIIEQKQKLSKILPRYQELT---KBEAMY--KLQAS 375
QY 156 LKQSSNSRKRTSAGTPVDRDNDGIPDSLEVEGYVDVKNKRTFLSPWISNTHK-KK 214
Db 376 LQKQDRDLILKGEYARFKSKDRDTHIHSIE-----ELKSS-----IQNLNLES 422
QY 215 GLTKVSSPKSTWSDPSYDPKVTGRIDKNVSPPEARPLVAAPVIVHVDMENIILSKN 274
Db 423 QLQMDRTSLRKQYSAID-----EIEELDLSINGPTKQL-----EFDSELIHLKQK 471
QY 275 EDGSTQNTDSETHIKNSTSTSTHSTSEVHGNAEVAHFFDGGVSGAGFSNSSTVAI 334
Db 472 LSESLDRKELRKEQKQLVLETLSDVNQNR-----NVNETMSRSLANGIINVKEI 525
QY 335 DHSISLAGE-----RTWARTMG-----LNTADTARLNANIRVYNTG 370
Db 526 TEKLSKSPESVFGTGLIKNDKYKCAEIVGNSLFIHVDTETATLWNLNLYRMKG 585
QY 371 TAPIYVLPITSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIALNAQKDFST 430
Db 586 GRVTF--IPLNRLSLDSVDFPSNTTQIQFTPLIKKIYEPFEKA-----VGHVFGKT 638
QY 431 PITWYVNOFLEKTKQLR---LDTQV-YGNLAT--YNFENGRVVDVTSKNWSEVLPI 484
Db 639 IVVKDLGGGLAKKHKLNATLDDGRADKRGVLTGGYLDQHKRTRLESKLNLSRSOH 598
QY 485 QETTARIIFNGKOLNVERIAAVNPS-----DPLETTKPD-MTLK 524
Db 699 KKILEELDFVRELNDIDTKIDQVNGIRKVSNDRESVLNIEVYRTSLNTKKNKXILE 758
QY 525 EALK-IAFENPENGNYQKQKIDTFDN----FDOQTQNTKNQLA-----E 568
Db 759 ESLNAILKLEKLNTRTFPAQEKLTNFENDLLOEFDSELSKEKERLESITKEISAHNK 818
QY 569 LNATN-----IYTVLDIKLNKKNILIRDKRFHYDRNNI-AYG-----A 607
Db 819 LNITSDALEGTTTID--SLNAELESKLIQENDLESKMGSEVGDATIFGQDELKELOLE 876
QY 608 DESVVKRAHREVINSST-----EGLLI-----NIDKDKIRKI 638
Db 877 KESVEKQENHNAVLEGTQVREBIESLIAETNNKKLEKANNQOQLLKKLDNFQKSEKT 936

QY 639 LSGYVIVEIEDTEGLKE-----VINDRYDM-----LN-----ISSLRQDGK 673
Db 937 MIKXTTLVTRRELOQRREIGLLPEDALPEDALSDITSDQLLQRLNDMTNMTISGLKNVNR 996
QY 674 -TFIDFKYNDK-----LPLYSNPYKNVAVAVTKENTIIINPSENG 714
Db 997 RAFENFKFNERKDLAERASELDESKDSIQDLVLKQKKNV-----DSIFQKYSNF 1052
QY 715 DT-----STNGIKKILIFSK 729
Db 1053 EAVFERLVPRGTAKLIHRK 1072

RESULT 7
SPOF SCHPO
ID SPOF SCHPO STANDARD; PRT; 1957 AA.
AC Q10411; Q9USE9;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sporulation-specific protein 15.
GN SPO15 OR SPAC1F3.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
ON NCBI_TaxID=4896;
RX [1]
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
RX MEDLINE=20107136; PubMed=10639340; Shimoda C.;
RA Ikemoto S., Nakamura T., Kubo M.,
RT "S. pombe sporulation-specific coiled-coil protein Spo15p is localized
to the spindle pole body and essential for its modification";
RL J. Cell Sci. 113:545-554 (2000).
RX [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicakart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Bortym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wadler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880 (2002).
RX [3]
RP SEQUENCE OF 705-871 FROM N.A.
RC STRAIN=968 h90;
RX MEDLINE=20223868; PubMed=10759889;
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
fission yeast cells by the use of a GFP-fusion genomic DNA library";
RL Genes Cells 5:169-190 (2000).

CC -!- FUNCTION: Has a role in the initiation of spore membrane
CC formation.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Spindle pole body.
CC -!- SIMILARITY: Belongs to the Mpc70 family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z70690; CAA94624.1; -;
CC EMBL: AB027811; BAA87115.1; -;
CC PIR: T38077; T38077.
CC GeneDB: Spombe; SPAC1F3.06c; -;
CC Sporulation; Coiled coil.
CC KW DOMAIN 199 785 COILED COIL (POTENTIAL).
CC FT DOMAIN 804 1235 COILED COIL (POTENTIAL).
CC FT DOMAIN 1320 1471 COILED COIL (POTENTIAL).
CC FT DOMAIN 1481 1723 COILED COIL (POTENTIAL).
CC SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;
CC
CC Query Match 4.5%; Score 168; DB 1; Length 1957;
CC Best Local Similarity 20.5%; Pred. No. 1.1;
CC Matches 178; Conservative 132; Mismatches 327; Indels 232; Gaps 40;
CC
QY 2 VKQENRLLESE-----SSQGLIGYFSDINQF--APMVVTSSTTGDLSPSSE 49
DB 129 VTQKNLLNELKQVRKLALEHENGILSLQSSNKKDKNTSSVTLTSEEDSVFQKK 188
QY 50 LENIPSENYQFSAIWS--GFIKV-----KKSDEYTFATSDADNHVTVMWVDQEVINKAS 101
DB 189 LTNMESNFSKQSEAYDLSQLLTVTEKLDKKEKDYKEDVSSIKASLAEEQASNKL 248
QY 102 NSNKIRLEK--GRLVQIKIYOENPTE---KGLDFKLYWTDSONKKEVISEDNLQPE 155
DB 249 RGEQERLEKLLVSNKTVSTLTOTENSRAECKTLQEL-----EKCAINEEDSKLEE 302
QY 156 LKQKSSN-----SRKRSSTAGPTVPDRDNDGIPDSLEVEGYTVDVQVNR 200
DB 303 LKHNVANYSDAIVHKOKLIEDLSTRISEFDN-----LKSERDTLSIKNEKLEKLRLNTI 356
QY 201 -TFUSPWISNIHEKGLTKYKSS-----PEKMWSTADSPYDFEK-----VTGRID----- 244
DB 357 GSLKDSRTSNLEEEEMVELKESNRTIHSQLTDAESKLSFQEENKSLKGSIDEXYQNLS 416
QY 245 -----KNVS--PEARHPLVAA-----YPIVHVDM----- 266
DB 417 SKRWVQVSSQLEAEARSLAHATGKLAETNSERDFQNKIKDFEKIEQDLRACLNSSN 476
QY 267 ---ENIILSKNEDOSTOND---SERTISKNVTS-----RHTSEVHGNAEVHASFPD 315
DB 477 ELKEKSAIDKQDQELNLRQIKQEKQKVSSTQSSLSQSLQDRDLNEKKKHEVYESQLNE 536
QY 316 IGGSVSAGFSNSG-----STVAIDHLSLAGEWTMAETMGLNTADTARLNANRYVNTG 370
DB 537 LKGLQTEINSEHLSQSLTLAEKAEAVATNNELSESK--NSLQTL-CNA----- 585
QY 371 TAPIYNVLPSSLVLGNQKQ-----TLATIKAKENQLSOLAPNNYFSPKNLAPALMAQKD 426
DB 586 ---FOEKLAQSVMLKENEQNFSSLDTSFKKLNESHOELENNHQTITKQL-----KD 634
QY 427 FSSPTITVNNYQFLELEKTKQLRLDTPQVYGNATYNFENGRVVDVTGNSWSEVLPOIQE 486
DB 635 TSS-----KIQQLQERANFEQKESITLSDENNDRTK-----LKLEE 672
QY 487 TTARIIFNGKOLNIVERRIAAVNPSDPLETTKPM-TLKEALKI-----AFGNEPKNLQ 541
DB 673 SNKSLIKQEDVDLSLEKNI-----QTLKEDLRKSEALRPSKLEAKNLREVIDNL- 722

QY 542 YQCKDIT-BFDENFDQTSQNIKNQLAELNATNIYTVLKDIKLNAGNLIIRDKRPHYDR 600
DB 723 -KQKHTELEAQRNDLHSSLSDAKNTWAILSSSELTXSSEDVKRLTANVETLTQDSK----- 776
QY 601 NNIAVGADESIVKAEHREVINSSTEGLLNIDKDIR-----KILSGYIVEIE-----DT 649
DB 777 -----AMQKSTSLVNSYQS--ISNLVHELDRDDHVMQSQNNLTLESSEKLTQDC 824
QY 650 EGLKE--VINDRYDMLNISSLRQDQKTFIDFKKYNDKLPYISNPNYKVNVAVTKEKT 706
DB 825 ENITQNMULDVQKLMKHVNVQESKV-SELKEVNGKLSLDLKNLLESSLNV-AISDNQ 882
QY 707 IINP-----SENGDT-----STNGIKKI 724
DB 883 ILTQLAELSKNYDSLEQESAQLNSGLKSL 911
RESULT 8
WAPA_BACSU STANDARD; PRT; 2334 AA.
AC Q07833;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Wall-associated protein precursor.
GN WAPA OR N17G OR BSU59230.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93302506; PubMed=8316082;
RA Foster S.J.;
RT "Molecular analysis of three major wall-associated proteins of
RT Bacillus subtilis 168: evidence for processing of the product of a
RT gene encoding a 258 kDa precursor two-domain ligand-binding
RT protein."
RL Mol. Microbiol. 8:299-310(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=95219088; PubMed=7704263;
RA Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;
RT "Cloning and sequencing of a 29 kb region of the Bacillus subtilis
RT genome containing the hut and wapa loci."
RL Microbiology 141:337-343(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=97124196; PubMed=8969509;
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT containing the lic and cel loci, and creation of a 177 kb contig
RT covering the gnt-sackY region."
RL Microbiology 142:3113-3123(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritze C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.I., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haeck J., Harwood C.K., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Huilo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigovic C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Neone D., O'Reilly R.P., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwoll G., Prescott A.M.,
RA Presecan E., Pujic P., Fumelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler K., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*,"
RL Nature 390:249-256(1997).
CC -!- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,
CC MOTILITY, SECRETION OR DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED
CC INTO THE MEDIUM.
CC -!- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE
CC 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE
CC C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED
CC MOTIF REPEATED 31 TIMES.
CC -!- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME
CC SIMILARITY TO THE REPEAT IN E. COLI RBS GROUP OF PROTEINS (RBSA-D).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L05634; AAA22883.1; -;
CC EMBL; D31856; BAA06556.1; -;
CC EMBL; D29985; BAA06260.1; -;
CC EMBL; D83026; BAA11683.1; -;
CC EMBL; Z99124; CAB15959.1; -;
CC PIR; S32920; S32920; wapa.
CC Subtilist; BG10797; wapa.
CC InterPro; IPR003305; CBM_CenC.
CC InterPro; IPR006530; YD.
CC Pfam; PF02018; CBM 4_9; 1.
CC Pfam; PF05593; RBS repeat; 14.
CC TIGRPFAM; TIGR01643; YD repeat 2x; 17.
CC Cell wall; Repeat; Signal; Complete proteome.
CC SIGNAL 1 228 OR 32 (POTENTIAL)
CC CHAIN 29 2334 WALL-ASSOCIATED PROTEIN.
CC DOMAIN 504 869 3 X 101 AA APPROXIMATE TANDEM REPEATS.
CC REPEAT 504 605 1-1.
CC REPEAT 636 736 1-2.
CC REPEAT 769 869 1-3.
CC DOMAIN 1021 2139 31 X 21 AA APPROXIMATE TANDEM REPEATS OF
CC X(4)-G-X(4)-[YF]-X-D-X(2)-G-X(4).
CC REPEAT 1021 1040 2-1.
CC REPEAT 1042 1061 2-2.
CC REPEAT 1063 1082 2-3.
CC REPEAT 1083 1102 2-4.
CC REPEAT 1109 1128 2-5.
CC REPEAT 1129 1148 2-6.
CC REPEAT 1150 1169 2-7.
CC REPEAT 1174 1193 2-8.
CC REPEAT 1199 1218 2-9.
CC REPEAT 1219 1238 2-10.
CC REPEAT 1646 1665 2-11.
CC REPEAT 1667 1686 2-12.
CC REPEAT 1690 1709 2-13.
CC REPEAT 1711 1730 2-14.

FT REPEAT 1732 1751 2-15.
FT REPEAT 1753 1772 2-16.
FT REPEAT 1795 1814 2-17.
FT REPEAT 1820 1839 2-18.
FT REPEAT 1840 1859 2-19.
FT REPEAT 1861 1880 2-20.
FT REPEAT 1887 1906 2-21.
FT REPEAT 1908 1927 2-22.
FT REPEAT 1929 1948 2-23.
FT REPEAT 1969 1982 2-24 (APPROXIMATE).
FT REPEAT 1983 2002 2-25.
FT REPEAT 2008 2027 2-26.
FT REPEAT 2028 2047 2-27.
FT REPEAT 2051 2070 2-28.
FT REPEAT 2071 2090 2-29.
FT REPEAT 2093 2112 2-30.
FT REPEAT 2120 2139 2-31.
SQ SEQUENCE 2334 AA; 258329 MW; B75138CCD279BAA3 CRC64;
Query Match 4.4%; Score 166.5; DB 1; Length 2334;
Best Local Similarity 18.2%; Pred. No. 1.7;
Matches 173; Conservative 126; Mismatches 336; Indels 317; Gaps 42;
Qy 1 EVQENRLLNESSSQGLLYGFFSLNFOA-----PMVVTSSSTGDLSPSSSELENIP 54
Db 269 EVERSQVSYKLEKNEBEGYLLHLTADENMLKOPERYVPSIDPSTSLSVSDTFVMSAYP 328
Qy 55 SENQYFQSAWSGFIK--VKKSDEYFATSADNHVTM----- 89
Db 329 ITNYSASSQKQDANLKAIVLKTGYD-KITGTYAFMKFNLLKPIQNMVTKATLKYVA 387
Qy 90 -----WVDDEVINK-----ASNSNKIRLEKGR--LQIKIQYOR 122
Db 388 HSYGTGKATGLWD--TVNSYDNAKVWNTKPAKSKNIGKADVHKQGWASVDVTAAVKS 444
Qy 123 ENTEKGLDFLKTWTDSONK---KEVSSDNLQLPKQKSSNSRKKRSTKSAGTPVP--- 176
Db 445 WNSGGANYGFKLH-TWNGSKYWKLLS-----SANSANKPIEVITYTIPKGN 491
Qy 177 -----DRNDGIPD-----SLE-VEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSP 223
Db 492 TPTIKAYHNGDSTGYFDISWKKVEGAKYKWIYNGKEYQAIASAGNV-----TSMSTKG 545
Qy 224 EK-WSIASDPYSDFEKVTGRID-KNVSPEARHPLVAAYPIVHVDMENILSKNE----- 275
Db 546 KKIWPTSAEIASKRYKL--HLDGKGAEALDP-----SPYKNSGGSYATSKNYWIGVSA 599
Qy 276 --DQSTONTDSETRTISKNTSTRTSTSEVHGNAEYHASFDFDGGSVSAG-----FS 325
Db 600 IFDQEGAMSAPAKPVPINVGKQAPSAKGYNNNGNA-TGYFDLSWKAIVSGATGYKYQVFN 658
Qy 326 NNSSTVAIDH--SLSLAGERTWAETMGLNTADTA-----RLNANRYVNTGYTA 372
Db 659 KGQFETLDLGNQTSWTKGKKIWIPTSAETKAGKYALHLKDGSGAELPINSPTGTYKNAGGD 718
Qy 373 PIYNVLPTTSLVLGKQNTLATIKAKENQLSQILAPNN-----YYPK-----NLA--- 417
Db 719 GAKRNTYFKIIAYNKDGEATASPAAPALPDIAAPKNTVGTLYTNTKSSQTVVNLWEK 778
Qy 418 -----PIALNAQKDFSSPTI-----TMNVTNQFLELEKY-----QLRL 450
Db 779 VQNAKGYKVINIYNGKEYQSPDVGADHWTQNKNIWPTSEIKAGSYKLHLDGKGELAL 838
Qy 451 DTQVYGN-----IATYN--PENGVRVVDVTSNWSSEV-L 491
Db 839 DPSPVNNANGNYKGVKQVSYFLVAYDANGETIPTAPFNPTFHEGAEFLGTEYVNIIDI 898
Qy 482 P--QIOETTARIIFNGKDLN-----VERIAAVNPSPDL-----FTT----- 517
Db 899 PSQQLNGATGNVIVNEEDLSIDQRGPGGLGSLRTYNSLDSSDLHFGQGWYDAETSVISTD 958
Qy 518 -----KPDWTLKEALKIAFG-FNEPENGNIQYQKDIETDFDFDQDQTSQNIQNLAE-L- 569


```
Db 959 GGAYVIDEDATTHRFKAKDGTGYPTGYVLELTETADQFIKTKVDQTNAYFNKGGKQLQ 1018
Qy 570 -----NATNIYTVLDKIKLNARONILIRDKRFHYDRN-----NIAVGADES 611
Db 1019 KVVVDGHNNATVYNDKQNLTAITDASGRKLTFTYDENGHVITSITGPKNKVKVTSYENDL 1078
Qy 612 VKE-----AAR--EVINSSTEGILL 629
Db 1079 LKKVTDGTVTSYDVSSEGLVQVQSANSTEAKPVFTFYQSGHRAKAINAKKETVY 1138
Qy 630 NIDKDI-----RKILSGY-----IVEIEDTEGLKEVINDRYDMLNI 665
Db 1139 SYDADKKTLLMTQPNRKYQVGYNEAGNPIQVDDAEGLKITNTKYEGNNV 1190

RESULT 9
ID_YJL3_YEAST STANDARD; PRT; 1803 AA.
AC P47024; P87192;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transposon Ty4 207.7 kDa hypothetical protein.
GN TY4B OR YJL113W OR J0780.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=97103775; PubMed=8948101;
RA Czepluch C., Kordes E., Pujol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including URA2 (5' end), TRK1, PPS2,
RT SPT10, GCD14, RPE1, PHO85, NCA3, ASF1, COT7, GZF3, two tRNA genes,
RT three remnant delta elements and a Ty4 transposon.";
RL Yeast 12:1471-1474 (1996).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z49389; CAA89409.1; -
CC SGD; S0003649; YJL113W.
CC InterPro; IPR001584; Rve.
CC InterPro; IPR001878; Znf_CCHC.
CC Pfam; PF00665; rve; 1.
CC SMART; SM00343; Znf_C2HC; 1.
CC Transposable element; Hypothetical protein.
SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;

Query Match 4.3%; Score 164; DB 1; Length 1803;
Best Local Similarity 19.5%; Pred.No.1.6;
Matches 162; Conservative 123; Mismatches 274; Indels 270; Gaps 41;

Qy 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSTGDSIPSSSELENIPSENQYF 60
Db 1078 QLKAKTN---HETSPFKSGSIG---TNVVFNTNNEISLTKGTSLPIKLESINHH--- 1128
Qy 61 QSALWSGFIKVKSDSEYFATSDADNHTVMVDQEVINKASNSKIRLEKGLYQIKIY 120
Db 1129 -----SNDY-----STNKVE-----KF 1140

Qy 121 QREN----PTEKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNRKRSTSGPTVP 176
Db 1141 EKENHHPPPIEDIVDM-----SQDTMESNCQDGNLKLKVTDKNVPTDNGTINVSPR-- 1193
Qy 177 DRNDGIPDLEVEGYTVDNVKNKRTPLSPWISNIHEKK-----GLTKY----- 219
```

```
Db 1194 -----LEQNEASGSPVQTVNKSATFLNKEFSLNKKRKHHRKNSUTSVELEDRDKR 1247
Qy 220 -----KSSPEKSTASDPYSDPEKVTGRIDKNVSPEARHPLVAAPIVHVDMENILSK 273
Db 1248 SKNRVKLIIDNNEVTSAPKIRAIYNEALSXPDLKERHEYKQAY---HKELQNLKDWK 1304
Qy 274 NEQSTQNTDSETRTISKN--TSTSTHTSEVHGNAEVHASPFDDIGGSVASGFSNSNST 331
Db 1305 VFDVDVKYRSE---IPDNLIVFTNTFTFKRNG---IYKARIVCRGDTQSPDTYSVIT 1358
Qy 332 VAIDHS-----LSLAGERT-WAETMGINTADT-ARLNANIRVYNTGTAPIYVLP--TTSL 383
Db 1359 ESLNHNHIKIFLAIANNRNMFMKTLIDINHAFLAKLEE-----IYIPHPDHRRC 1408
Qy 384 VLGNQTLATIKAKENQ-----LSQI-LAPNNTYP-----SKNLAFIALNAQKDFS 428
Db 1409 VVKLNKALYGLKSPKEMNDHLRQYLNGIGLKNSYTPGLYQTEDKXLM-IAVYDDCVI 1467
Qy 429 STPTIMYNOFL-ELEKTKQLR-----LTDQVYGNATYFENGCV-----469
Db 1468 AASNEQRLDEFINKLSNFELKITGLIDVDLTD-ILGMDLVYKRLGTIDLTLSFIN 1526
Qy 470 RVDTSNWSSEVLQIOETTARIIFNGKOLNVERRIAANVP-SDPLETTKPD-----MTL 523
Db 1527 RMD--KYNBELKKIRKSIPHMTYK-----IDPKKDLVQMSSEEFROGVKL 1573
Qy 524 KEALKIAGFNPENGNLQYQGDITFDFNFDOOTSONIKNQLAELNATNIYTVLDKIKL 583
Db 1574 QQLL-----GELNYV-RHKCRYDIEFAVKVVARLVNYPHERVFMVYIKIQL-- 1620
Qy 584 NAKNHLIRDKRFHYDRN-----NIAVGADES VVKEAHRE-----VI-----NS 622
Db 1621 -----VAYKDTIGIHYDRDCKKVKVIAITDASVGSYDAQSRIGTILVYGMNIFNYSNK 1675
Qy 623 STEGLLANIDKIRKILSGV-----IVEIEDT-----EGLK-----653
Db 1676 STNRCVSTSEALHAIYEGYADSETKVTULKEGBGNDNDIVMTDSKPAIQGLNRSYQ 1735

654 -----EVINDRYDMLNISLRQDGKFTI-----DFPKY 681
1736 PKERFTWIKTEIIKEIKESIKLLKITGKGNIADELLTKPVASADPKRF 1784

RESULT 10
ID_P115_MYCHR STANDARD; PRT; 979 AA.
AC P41508;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE P115 protein.
OS Mycoplasma hyorhinis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2100;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=91138990; PubMed=1825306;
RA Notarnicola S.M., McIntosh M.A., Wise K.S.;
RT "A Mycoplasma hyorhinis protein with sequence similarities to
RT nucleotide-binding enzymes.";
RL Gene 97:77-85 (1991).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- DOMAIN: Consists of two putative central coiled-coil regions
CC flanked by putative globular regions at the N- and C-termini.
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYCOPLASMA
CC P115.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; M34956; AAA25423.1; --
 DR PIR; JQ0894; JQ0894; ABC transporter.
 DR InterPro; IPR003433; GTP-binding_dom.
 DR InterPro; IPR005289; GTP-binding_dom.
 DR InterPro; IPR003405; SMC C.
 DR InterPro; IPR003395; SMC N.
 DR Pfam; PF02483; SMC C; 1.
 DR Pfam; PF02463; SMC N; 1.
 DR TIGRfam; TIGR00650; MG442; 2.
 KW ATP-binding; Coiled coil.
 NP_BIND 32 39 ATP (POTENTIAL).
 FT DOMAIN 169 224 COILED COIL (POTENTIAL).
 FT DOMAIN 231 400 COILED COIL (POTENTIAL).
 FT DOMAIN 569 821 COILED COIL (POTENTIAL).
 FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).
 SQ SEQUENCE 979 AA; 110566 MW; 30D51C56B56280F4 CRC64;

 Query Match 4.3%; Score 162.5; DB 1; Length 979;
 Best Local Similarity 18.9%; Pred. No. 0.81;
 Matches 175; Conservative 133; Mismatches 341; Indels 277; Gaps 40;

 QY 13 ESSSQGLGYFSDLNFPQPMV-----TSSTTGLSIPSSSELENIPSE 56
 DB 51 EQSAKQLRGLNDDVIFAGSKTVKPOEKAMVKLTFKNEAIAETKQIFTSILLKRGQGT 110
 QY 57 NQYFQS-----AIWGFHKYKSDYTFATSNADNHVMVDDQE--VINKASN 102
 DB 111 NEFYNDQPVRYKDKIKNLAVESG---ISKSLAISQGTISBIAETPQRKAVIEEAAAG 167
 QY 103 SNKIRLEKRLYQIKIQYQRENPTKGLDFKLYWTDSQNKKEVSDNQLQPELKQSSN 162
 DB 168 TSKYKLDKEAOKLRL-----TNDAD-KLQGAKELEKQVNSLDK-QASKAKIYLEK 219
 QY 163 SRKRRSTAGTPVDRD-----NDGIPDSLEVEGYTVDVK-NKRTFLSPWTSNIHEKKG 215
 DB 220 SKALESVEGLVINDLNFFNEKLNLTSLLEVEQQRNDLELNLIQTYESSISQTVH---- 275
 QY 216 LTKYKSSPEKWSASTDPYDFEKTGRID--KNVSP-----EAPHLVAAYPIVHV 264
 DB 276 ---FKTEVE-----SQIQTISKLNKLKNSLQALINQEAIEERKLIISGEIV-V 322
 QY 265 DMENII--LSKN-EDOSTQNTDSEITISKNSTSTSTH----- 300
 DB 323 DQKTKIEIKKQVESLKIQINASKQREIELDQQLTRLNKANSKLQENDINKIEIGVLE 382
 QY 301 --SEVHGNAEIVHAFPDIGGSVAGSFGNSSTVAIDHSLSLAGERTWAE---TWGLNTA 355
 DB 383 KKSAAAANILKQFQENKSFSLKGIKTIKNSFLDFGVIGLASFLKVESEFSLAIEIV 442
 QY 356 DTLRLN-----ANIRYVNTGTAPIYVNLPTTSL-----VLG 386
 DB 443 LGAALNIVMKTSEDVLOAIDFLKKNLSGKATFPLTSIKEREVEDHLLVLKGQKGLG 502
 QY 387 KNOTLATIKAKNO-----LSQILAPNNYPSKNLAPI-----ALN 422
 DB 503 VAKELIEFDQFNKFLGFLGLNLIIVDNDVNANRIKILDPKHYTIYVLSGLDFRPGGTIT 562
 QY 423 AQKDFSTPTIMYNQPLEKTKQLRLDQ-----VYGNIATVNPENGRVVD 472
 DB 563 GGSKLERTSI-LNVDIKIK-EHTNWLKFAEDQIHLKIKQOTIYNIETVNSTIQOVKIE 620
 QY 473 TGS-----NWSEVLPOIQE--TTARIIFNGKDLNLVERRIAAV----- 508
 DB 621 ANSINSKLNILNEELNKLNASIEFKQEQQEDQESLNSFDSKLN-IEKQISTLATIELN 679
 QY 509 NPSDPL-----ETTKEDMTLK-----EALKIATGFPENPNQYQCK 545
 DB 680 SKDRLTNLISEQKGETKKQELDAKRLKNTQHSDSITEQNRAKFLVEQNKRLSEHK 739
 QY 546 DITE-----FDNFQDQTS--QNIKNQALNATNIYVLDKIKLNKQNLILDKRF 596

Db 740 LTLTAAAEQYSLDLTIEQARHFVDSLKKEKELGNVNLNLEAITFEFVNQR-----YOEKKQ 795
 QY 597 HYDRNNIAGADESVVKEAHRREVINSSTEGLLNIDKD-----IRKILSGYIIVEIDTEG 651
 Db 796 YIBELITAKSKIIEAISLDLDKIINKTTE--IVNLVNNERNMVFQKFGGKAEIHFT-- 851
 QY 652 LKEVINDRYMLNIS-----SLRQOKTKFIDPKKND-----KLRLYI- 689
 Db 852 -----DKNDILNSGVEISAQPPKTKIKRLFSGGEKAIISLFLFAIKARPIPLCIL 905
 QY 690 -----SNPNVKNVYAVTKENT 706
 Db 906 DEVEALDESNNVRYVEFLKLLKENT 931

 RESULT 11
 BND_CLOBO STANDARD; PRT; 1276 AA.
 AC P19321; CLOBO
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)
 DE (Bontoxilin D).
 GN BOTD.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BYD/-3;
 RX MEDLINE=91016853; PubMed=2216736;
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
 RA Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;
 RT "Nucleotide sequence of the gene encoding Clostridium botulinum
 RT neurotoxin type D";
 RL Nucleic Acids Res. 18:5556-5556(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CB16;
 RX MEDLINE=93042276; PubMed=1420572;
 RA Surugawa H., Ohya T., Watanabe T., Inoue K.;
 RT "The complete amino acid sequence of the Clostridium botulinum type D
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RT phase d-16 phi genome";
 RL J. Vet. Med. Sci. 54:905-913(1992).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC STRAIN=D-SA, and D-1873;
 RX MEDLINE=89339741; PubMed=2668193;
 RA Morishita K., Syuto B., Kubo S., Oguma K.;
 RT "Molecular diversity of neurotoxins from Clostridium botulinum type D
 RT strains";
 RL Infect. Immun. 57:2886-2891(1989).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94230352; PubMed=8175689;
 RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
 RA Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
 RT E botulinum neurotoxins and tetanus toxin";
 RL J. Biol. Chem. 269:12764-12772(1994).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-LEU-61 BOND OF
 CC SYNAPTOSOMAL VESICLE-1 AND -2.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No

CC detected action on small molecule substrates.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity). (L) and a
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H). The light chain has the pharmacological activity,
CC while the N- and C-terminal of the heavy chain mediate channel
CC formation and toxin binding, respectively.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -1- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
CC BACTERIOPHAGE.
CC -1- SIMILARITY: Belongs to peptidase family M27.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54254; CAA38175.1; -;
CC EMBL; S49407; AAB24244.1; -;
CC PIR; S11455; S11455.
CC HSP; P10845; 3BTA.
CC MEROPS; M27.002; -;
CC InterPro; IPR008985; ConA like lec_g1.
CC InterPro; IPR002160; Kunitz legume.
CC InterPro; IPR006025; Pept_M2n_B3.
CC InterPro; IPR000395; peptidase_M27.
CC Pfam; PF01742; Peptidase_M27; 1.
CC PRINTS; PR00760; BONTOKILYSIN.
CC ProDom; PD001963; Bontokilysin; 1.
CC PROSITE; PS00142; ZINC PROTEASE; 1.
CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
CC CHAIN 1 442 BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.
CC METAL 443 1276 BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.
CC ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 230 230 BY SIMILARITY.
CC METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
CC DISULFID 437 450 INTERCHAIN (PROBABLE).
CC VARIANT 15 16 ND -> PV (IN STRAIN D-1873).
CC VARIANT 17 18 ND -> Q (IN STRAIN D-1873).
CC VARIANT 452 452 K -> Q (IN STRAIN D-SA).
CC VARIANT 457 457 R -> T (IN STRAIN D-SA).
CC VARIANT 457 457 R -> F (IN STRAIN D-1873).
CC VARIANT 462 462 A -> D (IN STRAIN D-1873).
CC VARIANT 489 489 K -> N (IN STRAIN CB16).
CC VARIANT 644 644 N -> K (IN STRAIN CB16).
CC VARIANT 1122 1122 Q -> R (IN STRAIN CB16).
CC SEQUENCE 1276 AA; 146871 MW; C1E50F46C8233E2 CRC64;
Query Match 4.3%; Score 161; DB 1; Length 1276;
Best Local Similarity 20.6%; Pred. No. 1.4;
Matches 164; Conservative 117; Mismatches 291; Indels 224; Gaps 41;
QY 44 SIPSELENIENQYFQSAWMSGFVKVKSDEYTPATSDNHHVMMVDDQEVINKASNS 103
DB 310 TIPSSWTSNI---DKY--KKLFSEKYNFDKNTGNTGVNIDKNSLYSDLTNMSVWVS 364
QY 104 NKIRLEK-----GRLYQIKIQYQRENTEKGLDFKLYWTSQNKKEV 145
DB 365 SQYVNVKTHYFSRHYLPVPANILDDNIYTIIRDGF--NLTKGFNIE----- 409
QY 146 ISSNLQL-PELKOKSSNRKRSKTSAGTPVDRDNGIPDSLEVEGYVDVKNKRTFLS 204
DB 410 NSGQNIERNPALQKLSSVVDLTFKVCLRTKNSRD-----DSTCIKVNKRL--- 458
QY 205 PWISN-----IHEKGLTKYKSPKWKSTASDPYS-DPEKVTGRIDKQVSPKARHPLV 256
DB 459 FVADKDSISQEIFENKLIITD-ETNVQNY---SDKFSILDESILDGQVP--INPEIVDPL- 511

QY 257 AAVPIVHVDMENILSKNE-----DOSTQNTD-----SETRTISKNTS--TSRTHTSEV 303
DB 512 ----LPNVNNEPLNLPGEEIVFYDDITKYVDYLSYVYLESQKLSNNVENITLTTVVEEA 567
QY 304 HGNAEVASPF-----DIGGSVAGSNSNSSTVAIDHSLSLAGERTWAETM----- 350
DB 568 LGYSNKIYTFPLSLAEKVNKGVOAGLFLNANWEVDEFTTNIMMKDITLDKISVSUWIPIY 627
QY 351 --GLNTADTA-RLNANIRVYVNTG-----TAPIYVNLPTTSLVLGKQNTLATIK 395
DB 628 IGPALNIGSALRGNFQAATAGVAFLLSGFFPEFIPALGVFTFYSSIQEREKIINTIE 687
QY 396 -----AKENQLSQILAPNNY-----YPSKNLAPIALNAQKDFSSPITM 434
DB 688 NCLQRVQRWRKDSYQWVSNLWLSRITTFQFNHINYQYDS-----LSYQADAIAKADIL 740
QY 435 NYNQFLEK-----TKOLRLDTDQVYGN-----ATYFNGRVRVDTGSN 476
DB 741 EYKISSGDKENIKSQVENLKNLSVDVKISEAMNINKFIRECSVTYLYFKN----- 790
QY 477 WSEVLPQIQETTAIIIFNGKD--LNL-----VERRIAAVNPSDPLETTKPDMTL 523
DB 791 --MLPKVIDELNKFDLRTKTLEINLIDSHNIIIVGEVDRLLKAKVNES--FENTMP--- 841
QY 524 KEALKIAGF-PNEENGNLQYQGDITEFDNFDOQTS-----ONIKNOLAELNATNI-YTV 577
DB 842 -----FNIFSYTNNSLL---KDIINEYFNSINDSKILSLQNKKALVDTSGYNAEVRV 891
QY 578 LDKIKLNAKNNILIRKRFHYDRNNIAGVADSVKVEAHREVINSSTEGILLNIDKDIRK 637
DB 892 GDNVQLN---TIYTNDFKLSSSGDKIIVLNNNLIYSA---IYENSVSFWIKISKDLTN 945
QY 638 ILSGYIV--ETEDTEGJKEVINDRYDMNLISLRQDGKTFIDKYNKDPLV---ISNP 692
DB 946 SHNEYTTINSIEQNSGKWLCLIRNG---NIEWILQDVN-----RKYKSLIFDISELSHT 996
QY 693 NYKVNVAVTKEHTII 708
DB 997 GYTNKWFPTITNNIM 1012
RESULT 12
RBPI_PLAVB
ID RBPI_PLAVB STANDARD; PRT; 2869 AA.
AC Q00798;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reticulocyte binding protein 1 precursor.
GN RBPI.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.,
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites."
RL Cell 69:1213-1226 (1992).
CC -1- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to
CC human reticulocyte cells.
CC -1- SUBUNIT: Homodimer (potential).
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; M88097; AAA29743.1; --
KW Malaria; Receptor; Signal; Transmembrane.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 2869 RETICULOCYTE BINDING PROTEIN 1.
FT DOMAIN 18 2807 EXTRACELLULAR.
FT TRANSMEM 2808 2826 POTENTIAL.
FT DOMAIN 2827 2869 CYTOPLASMIC.
FT SITE 1030 1032 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 2599 2601 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 2869 AA; 330213 MW; B9DBB442205EBCFF CRC64;

Query Match 4.2%; Score 159.5; DB 1; Length 2869;
Best Local Similarity 18.4%; Pred. No. 4.8;
Matches 146; Conservative 142; Mismatches 297; Indels 209; Gaps 38;

QY 10 NESSESSQGLGYFSD-----INFQAPMVVTSSTTGDLSIPSELENIENISENQYQS 62
DB 1944 NSMYESMVLTYFLSDEAKISSGMEFNAEMKSNFKTDLLEIFSV-----ISNSNEL--- 1996

QY 63 AIWSGFIKVKSDYTFATSDHNVHVMVDDEIVKASNNKIRLEKGRLLQIKIQYOR 122
DB 1997 -----LKIEQ-----DSNDVIQKERESQLAKDATDIYV-ILKKN 2032

QY 123 ENPTEKGLDFKLYWTDSQNKKEVISSDNLQLPELKOKSSNRKSTAGTVPDRDNDG 182
DB 2033 E-----FNEKLEAKNKEEVVS-----EKVREALKRLS----- 2060

QY 183 IPSLEVEGYVDVKKRTFLSPWISNIHE-----KKGLTKYSSPEKWTASDPYSDREK 238
DB 2061 -----QVEGRCHFEHFRLL-----DNTLEENLKQWVIYR---DKKSERSGQLQEMEN 2108

QY 239 VTGRIDKNVSPEARHPLVAAPIVHVDMENIL-----SKNEQOSTQNTDSETRTISKNTS 294
DB 2109 EMNTYSNSIT-----QLEGIVVSAGESKEDIEKLEERSNEENRNISEKIS 2152

QY 295 TSTRHTSEVHGNAEVHAFSDIGSVSAGSNNSSSTVALDHSLSL-----ACERTWAETM 350
DB 2153 TIDSKVIEMASTID---ELYKLGKNCQA-----HWISLIYSTANWTKTSKKLI 2196

QY 351 GLNATADTARLANIRYVNTGTAPYVLPVTTSLVGLKNQTLATIKAKENQSLAPNNY 410
DB 2197 MIN-KEKENTKCDVYIKDNSSSTGIVETLKGFGYKLTFFS-----ASEIVQNADT 2248

QY 411 YPSKNLA---PIALNAOKDFSSPTITWVYNOFLEL-----EKTQQLRLTDQV 455
DB 2249 Y-SYNFAKHEKESUNAIRDIKKELYLFHONSDISIVEGGVONMLALYKLEKREMDL 2307

QY 456 YGNATYVNFENGVRVDTGNSWSEVLPOIQTETARIIFNGKDLNVLERRIAAVNPSPDPLE 515
DB 2308 YRNISETKLKQMEHSTDVFXPMIELHKGMMET-----NNKSLLKXKXKLSVN--DHMH 2359

QY 516 TTKPDWTLKALKI-----AFGFNEPNGLQYQKGDITEFDNF--DOOTSQNTKQLAEL 569
DB 2360 SMEAEM-INKGLKYPSPVQNNINNIYGVIAEVKLTLEIDRDYGDNYQIVVEEHKKQFSIL 2418

QY 570 -NATNIYVLDKIKLNAKMNILLRDKRPHY---DRNNIAGADESVVKEAHREVINSST- 624
DB 2419 IDRTN--ALMDDIBIFKEN-----NYNLMVNTETTHRVNDVIEKTLKLVQAKTE 2468

QY 625 -EGLLLNI---DKDIRKILSGYVIEIEDTEGLKE-----VINDRYDM-----LNTSSLR 669
DB 2469 YEQILENIKQNDMLQNIIFLKVKYSIIEYFENVKXKESILNDLYEQERLLKIGELHDEIK 2528

QY 670 QDGKTFIDPKYNDKLPYISN-----PNYKVVYVYKENTLIINPS-----ENGDTLS 717
DB 2529 RNVETUSSYSDIQKEMMSKLLKESKMMNY--TSIYELEREANEINRAKQIKDDTTI 2587

QY 718 TNGIKKILIFSKKG 731
DB 2588 LNSVLEAAI-QKRG 2600

RESULT 13

S155_YEAST
ID S155_YEAST STANDARD; PRT; 1000 AA.
AC P43612;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S174-associating protein SAPI55.
GN SAPI55 OR YFR040W
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96220458; PubMed=8649382;
RA Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA Arndt K.T.,
RT "The SAP, a new family of proteins, associate and function positively
RT with the S174 phosphatase.";
RL Mol. Cell. Biol. 16:2744-2755(1996).
RN [2]
RP SEQUENCE OF 98-1000 FROM N.A.
RX STRAIN=S288C / AB972;
MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 98-1000 FROM N.A.
RX STRAIN=S288C / AB972;
MEDLINE=96287652; PubMed=8686379;
RA Eki T., Naitou M., Hagiwara H., Shibata T., Hanaoka F., Murakami Y.,
RA Sasanuma M., Tsuchiya Y., Shibata T.,
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of
RT chromosome VI from Saccharomyces cerevisiae.";
RL Yeast 12:149-167(1996).
CC -!- FUNCTION: Associates with the S174 phosphatase in a cell cycle
CC dependent manner. May be directly or indirectly involved in
CC S174-dependent functions in budding and in normal G1 cyclin
CC expression.
CC -!- PTM: Hyperphosphorylated in the absence of S174.
CC -!- SIMILARITY: Belongs to the SAPS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U50560; AAC49303.1; --
DR EMBL; D50617; BAA09279.1; --
DR GenOnline; 140194; --
DR SGD; S0001936; SAPI55.
DR InterPro; IPR007587; SAPS.
DR Pfam; PF04499; SAPS.1.
KW Phosphorylation; Cell cycle.
FT CONFLICT 663 663
FT CONFLICT 674 689
N -> T (IN REF. 2 AND 3).
TYSKSNYMIRDCFQNN -> DLFKIKLYDTRIVSKIM (IN
REF. 2 AND 3).
DY -> EL (IN REF. 2 AND 3).
RYSSN -> PDIOVI (IN REF. 2 AND 3).
FT CONFLICT 814 815
FT CONFLICT 818 822
FT SEQUENCE 1000 AA; 114880 MW; AD4BF0BC7C588D5 CRC64;

Query Match 4.2%; Score 159; DB 1; Length 1000;
Best Local Similarity 19.4%; Pred. No. 1.2;
Matches 172; Conservative 128; Mismatches 324; Indels 262; Gaps 42;

QY 1 EVKQENRLNESSSQGLGYFVS--DL-----NFQAPMVVTSSTTGD-----LSIPS 47

Db 90 EKEEERSNNSSESSTSTSSGSGTSTKTDLEEDISNATAPMMVTTKNLNDNGSIERMLVET 149
Qy 48 SELENIPSENQYFQSAINSGFIKVKSDYVATATSADNHNVTMMVDD-CEVINKASNSNKI 106
Db 150 EILNELSRQNKLLDPIFCGPFDDKINKV-----NMVELVQLMCEKSKIKATV 203
Qy 107 RLEKGLYQIKIQYQRENTEKGLDFKLYWTSQNKKEVISSDNIQLPELKQKSNRKK 166
Db 204 DLNN-----LIDYQBOQLD-----DSSQEDVYVESDTEBEKEDDNNNNKK 247
Qy 167 R-----STAGPVPDRDNDGIPDSLEVGIVYDVKNKTEFLSP-----WISNIHEKKG 215
Db 248 RRRGSSFGNDNNDDDDANEDDESAYL-----TKATIISEIFSDIMWIS-----ES 299
Qy 216 LTKYSSPEK-WSTADDPYSDPEKVTGRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKN 274
Db 300 LVNQSYLNKWSIIINQ-----NFNSE-NSPLV-----PIFLKINQNLILLR- 341
Qy 275 EDQ-----STQWTDSET-----RTISKNT 293
Db 342 QDQYLNPIFRTSFVDMKLVHVDISLMDFFKLISTDKIESPTGIIELVYQNLISKCL 401
Qy 294 STSRTHTEVHGNAEVAHSPFDIGSGVAGFNSNSSTVAI-DHSLS--LAGERTWAETM 350
Db 402 SPLNNKESP-----ADIQACVGDFLKALIALSANAPLDDISIGPNSLTRQLASPEIAKLV 457
Qy 351 GLNTADTARLANIRVNTGTABIYVNLPTSLVGLKQNTLATIKAKENQLSQILAPNNY 410
Db 458 DI-----MINQGAALNT-----TVSIVI-----ELIRKNSDYDQVLLTTT 498
Qy 411 YPSKNLAPIALNAQKDFSSPTIMVYNQFLEKTKQLRLTDQVYGNIAVTFNGVR 470
Db 499 HPSNRDPVYGLLRKFSNHLSDFFQIILDIENDANIPLHENQLHEKPKLGFERFKV- 557
Qy 471 VDTGSNWEVLPOIQTETARIIFNGKDLNLVERRIAANVP--SDPLETTKPMTL--KEA 526
Db 558 -----VELIAELLHCNMLGNLSKRAERARRDKVRSQLSHLLQDALNDLSIEKEQ 610
Qy 527 LKIAFG-----FNEPNCNLOVQGGDI-----TEFDFNFD-----QOTSNIKNOLAE 568
Db 611 LKTKHSTRTDHLKNNKINDNDNDDSDYGEIDESEFPIYINMKQNKILRTDP 670
Qy 569 LNATNIYVLDKIKLNAKMILIRDKRFHYDRNNAIAGD-----BSVVKAEHREVIN-- 621
Db 671 TVGT-----YKSNYMIRD-----CFQNNELFTHPWNFWNFVFDIIQQIFNGR 716
Qy 622 ---SSTGLLNLID--KDIRKILSGYIVETE---DTEGLKEVINDRYDMLNISS---LR- 669
Db 717 MDFSNGFLVSLFNLSKSYQFTMDIVISDEKGTDSVRSFVIRDPNFDKITTDFILRG 776
Qy 670 -QDGTETIDFKKN-----DKLPLYISN-----PNYKVNYY 699
Db 777 YQDSYKFEYLRKMLGYMGHVLIAEWEVKFSKLYKVDYISRYSSNLTQTEWQYSEVL 836
Qy 700 AVTK-----ENTIIINPSENGDTSTN 719
Db 837 NETRMWYSKILGGSGYIDDDNGNIIPOLPNTTVLTP--NGDASNN 880

RESULT 14

FAT1_SCHPO
ID -FAT1_SCHPO STANDARD; PRT: 1385 AA.
AC O13735; QOUT0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Actin interacting protein 3 homolog.
GN FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;
RN [1]
RC SEQUENCE FROM N.A.
RP MEDLINE=21848401; PubMed=11859360;
RX STRAIN=972;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Seguros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gelles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moutle S., Munnall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford J., Rutter S., Saunders R., Squares S., Stevens K.,
RA Skelton J., Simmons M., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Taylor K., Volkart G., Aert R., Robben J., Garzon A., Thode G.,
RA Woodward J., Volckaert E., Rieger M., Schaefer M., Mueller-Auer H.,
RA Weltjens I., Vansteede E., Fritsch C., Holzer B., Moestl D., Hilbert H.,
RA Gabel C., Fuchs M., Fritsch C., Holzer B., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC [2]
RP IDENTIFICATION, AND GENE NAME.
RX MEDLINE=20143585; PubMed=10679021;
RA Jin H., Amberg D.C.,
RT "The secretory pathway mediates localization of the cell polarity
regulator Alp3/Bud6p."
RL Mol. Biol. Cell 11:1647-661(2000).
CC -!- FUNCTION: Involved in the organization and/or function of the
actin cytoskeleton.
CC -!- SIMILARITY: TO YEAST BUD6.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; 297208; CAB10112.1; .
DR EMBL; AL109770; CAB52420.1; .
DR GeneDB SPombe; SPAC15A10.16; .
DR GO; GO:0005884; C:actin filament; ISS.
DR GO; GO:0016043; P:cell organization and biogenesis; ISS.
DR InterPro; IPR005613; AIP3.
DR Pfam; PF03915; AIP3; 1.
DR Coiled coil; Cytoskeleton.
KW DOMAIN 1009 1096
FT COILED COIL (POTENTIAL).
SQ SEQUENCE 1385 AA; 154325 MW; ABB3D40CC4FF7537 CRC64;

Query Match

Best Local Similarity 4.2%; Score 159; DB 1; Length 1385;
Matches 147; Conservative 125; Mismatches 293; Indels 202; Gaps 33;
QY 26 DLNQAPMVVTSSTGDL-----SIPSSLELTPSENQYFQSAIWS-----GPIKVKKS 74
Db 326 DRKFSPLKTRTFTSLTKSLDPTGTSLSXSPSLRKSFS--SFVQKDVYGRSRLSRISQANRS 383
QY 75 DEYTFATSADNHNVTMMVDDQSVINKAS-----NSN 104
Db 384 NVFPGATD-----VTRSVSDRILSSINDGEVAPLPQRSTISSPNSPLSATVLPST 439

105 KIRLEKGRLYQIKIOVQREN-----PTEKGLDFKL-----YWTDSQNK 143
 440 PILLRGRSSTLVNKKQFNADDGSLPNSIRTEYAAAPKLEDIADVEVDATISOR 499
 144 EVI-----SSNLQLPFLKQKSS---NSRKRKRSAGFTVPDRNDGI-----PDSL 187
 500 ELLERQIQKAESEDTSEISLQGLSLPQVSSTQCEIQPSSVPEAAASNAIEAKEPAVT 559
 188 EVEGTVDVGNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSFEKVTGRIDKNV 247
 560 ALESIT-----ERKEAPVISSEKIESGTSISTDTKGLANFENDSELEBLIQON- 613
 248 SPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRITISKNTSRTSRTHTEVHGNA 307
 614 NAEQDEPSYKFKHYEYSSEE-----SGSEDEPKSEKOTKGYVISNDDSTQVEEDSEDKSTP 669
 308 EVHASFFDIGGSVAGFNSNSSTVAID-----HSLSLAGERTWAEWTGLNTADTARLNAN 363
 670 NTGA-----SAKLINDPSSTITVSDVPKPKPASPEITEPPSSALVSATSPPTNVP 720
 364 I--RYVNTGTAPINVLPTTSVLGKNOTLATIKAKENQLSQILAPNNYPSKMLAPIAL 421
 721 IPVBAVHLSTA--FSTAPVSTIV-----SNISPLP*VAP--P 753
 422 NAKQDFSSPTIM--NYNQFLEKLT--KQLRLTDQVYGNIAIYNFENGVRVDTGSNW 477
 754 NVSGSPSETPIKPEKVPVSGTEKALPKPLGVDTERY--FLRYNNQTRKVRKVESPLN 811
 478 SEVLPOIQETTARI--FNGK--DLNIVERRIAAVNPSDPLETTKPDMT--LKEALKIAPGFN 534
 812 ANELGELFNVVKISFGSDSYELNIED-----POTKISYLLEDLSLKYSLVSPWFK 864
 535 EPNGLQ--YOGKIDITFDFNFDOQTQNTQNKQIAELNANIVT-VLDKTKLNAKMIL 590
 865 EQDANKREDFHSGEVSATQHSQAQNTLDDHVTNTTHESPSSAPTEILERLKA----- 917
 591 IRDKRPHYDRNNIAVGADESUVKEAHREVINSSTGLLNDKDKIRKILSGVIV-EIEDT 649
 918 -----IEQNIST-----NHTNDSAAKSEDSHKLANNFSPDSIDH 954
 650 EGLKEVINDRYDMLNISSLRQDKTFI-----DFKK-----YNDX 684
 955 KFYQVQKMQ--LELASLKQISAAFPTRIPLKIRKDKKEINAFNEK 998

RESULT 15
 CAGA_HELPJ
 ID CAGA_HELPJ STANDARD; PRT; 1167 AA.
 AC Q9ZLT1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytotoxicity associated immunodominant antigen (120 kDa protein)
 DE (CAG pathogenicity island protein 26).
 GN CAGA OR CAI OR CAG26 OR JHP0495.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]_TaxID=85963;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923692;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.,
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori."
 RL Nature 397:176-180(1999).
 CC -! FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,
 CC OR FUNCTION OF THE CYTOTOXIN.
 CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001483; AAD06073.1; -
 DR PIR: B71924; B71924.
 DR InterPro: IPR005169; Caga.
 DR InterPro: IPR004355; IVSec_caga.
 DR Pfam: PF03507; Caga; 1.
 DR PRINTS: PR01553; TYPE4S4CAGA.
 KW Antigen; Complete proteome.
 FT DOMAIN 245 249 POLY-THR.
 FT DOMAIN 882 889 POLY-ASN.
 SQ SEQUENCE 1167 AA; 129729 MW; FDS8E6B81CEBD0F2 CRC64;
 Query Match 4.2%; Score 158.5; DB 1; Length 1167;
 Best Local Similarity 19.7%; Pred. No. 1.6;
 Matches 159; Conservative 149; Mismatches 308; Indels 187; Gaps 41;
 QY 3 KOENRLNISESSQGLL---GYFSDJ-NF-----QAPMVVSTTGDLSIPSELE- 51
 DB 480 KADKALREAKTTLOGSLKHGDMFVDYGNFYKYNASKSPDKGVGATNG-----VSHLEA 534
 QY 52 -----NIPSENQYFQSAIWSGFIKVKKSDEYTPATSAADNHVTMWDD-----QEVIN 98
 DB 535 GFESKVAVFNLPNLNLAITSV---VRQLEDKLIAKGLSPQEAANKLVKDFLSSNKELVG 590
 QY 99 KASNSKIRLE---KRLYQIKIOVQRENPTKGLDFKLYWTDSDQNKKEVIVSDNQLPFE 155
 DB 591 KALNFNKAVAEAKNTGNYDEVK---QAQKDLKSL-----KKRLEKDVAK--N 635
 QY 156 LKQSSNSRKRKTSAGTPVDR-----DNDGIPDSLEVEGYTVDVKNKRTFLSPWISNI 210
 DB 636 LESKSGNKKVKAESQANSQKQEIFALINKEANDARAI--AYAQNLIKIGKRELSKLENI 694
 QY 211 HEKKGITKYKSPKRWSTASDPYSDEK-----VTGRI--DKNVSPERHPLVAAYPIV 262
 DB 695 N--KDLKFSKSPDEFNGKN--KDFSKAEETLAKLKGSVKDLGINPEWIS----- 741
 QY 263 HVDMENILSKNEDQSTQNTDSETRITISKNTSRTSRTHTEVHGNAEVHASFFDIGGSVA 322
 DB 742 --KVENLNAALNEFKNGKDKSKVTOAK--SLENSIKDVIINQKITDKVNLNCAVSV 797
 QY 323 GFNSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARLNANIRYVNTG-TAPIYVLPPT 381
 DB 798 AKATGDFSGV-----EQALADLKNFSKQLAQQAQKQEDNFTKNSALYQ----- 842
 QY 382 SLVLGKNTLATIKAKENQLSQILAPNNYPSKMLAPIALNAQKDFSSFTPIWVYQFLE 441
 DB 843 SVKNGVNGTIV-----GNGLSKAEAT--LSKNFSDI-----KKELNKLGKFNNNNNG 890
 QY 442 LEKTKQLRLTDQVYGNIAIYNFENGVRVDTGSNWSVLPQIQETTARIIFNGKDLNV 501
 DB 891 LENS-----TEPIYTOVA---KKVKAKIDRLDQIASGLGDVGQAASFLKHKHDKVDL 940
 QY 502 ERRIAAVNSPDPLETTKPDV---TLKEALKIATFAGNEPENGNYOYOGKIDITEFDNFDOQ 557
 DB 941 SKVGLSAN-HEPIYATIDDLGGFPFLKRHKV-----DDLSKVLGSLREKQ 984
 QY 558 TSNQIKN---QLAELNAT---NIYTVLDKIKLNKXNIIIRDKRPHYDRNNIAVGADES 611
 DB 985 LTQKIDNLNCAVSEAKASHFDNLDQMDIKLQSTKQNV-----NLVYESAKKV 1033
 QY 612 VKSAHREVINSSTEG---LLNLI--DKDIRKILSGYIVIEDTEGLKEVINDRYDMLNISS 667
 DB 1034 PTLISAKLDNYATNSHTRINSNVKNGTINEKATGMLTQ--KNSWLK--LVNDKIVAHNVGS 1091
 QY 668 LRQDKGTTFIDF-----KYNL--KLPLYISNP-----NKNVNVAVTKENTIIIPS 711

Db 1092 APLSAYDKIGFNQKNMKDYSDSFKSTRLSNAVKDIKSGFVQFLTNIFSMGYSIMKASV 1151

Qy 712 ENGDSTNGIKKILIFSKGYE 733

Db 1152 EHGKNTN-----TKGGFQ 1165

Search completed: May 3, 2004, 19:36:51
Job time : 13.2539 secs

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 34.6261 Seconds

(without alignments)
6697.424 Million cell updates/sec

Title: US-09-848-909A-8

Perfect score: 3773

Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKLIFSKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL_25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriaph.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3695	97.9	743	2 Q937W2	Q937W2 bacillus an
2	3695	97.9	743	2 Q937W3	Q937W3 bacillus an
3	883.5	23.4	876	2 C32739	C32739 clostridium
4	883.5	23.4	879	2 O06498	O06498 clostridium
5	881.5	23.4	876	2 Q9KH41	Q9KH41 clostridium
6	877	23.2	875	2 O46221	O46221 clostridium
7	819	21.7	721	2 O81W1	O81W1 clostridium
8	734	19.5	775	2 O844J8	O844J8 bacillus th
9	286.5	7.9	225	2 O8KXK2	O8KXK2 bacillus an
10	220.5	5.8	1881	16 Q8RKG2	Q8RKG2 fusobacteri
11	217.5	5.8	3468	5 Q81104	Q81104 plasmodium
12	205	5.4	4688	16 Q9PQ08	Q9PQ08 ureaplasma
13	202.5	5.4	1904	5 Q8IKS2	Q8IKS2 plasmodium
14	199.5	5.3	2849	5 Q8IHV4	Q8IHV4 plasmodium
15	194	5.1	1404	2 Q8RJN9	Q8RJN9 mycoplasma
16	191.5	5.1	675	16 Q897H6	Q897H6 clostridium

17	191.5	5.1	1301	5 Q8WSK5	Q8WSK5 plasmodium
18	191.5	5.1	2940	5 Q8IHP9	Q8IHP9 plasmodium
19	191	5.1	3248	5 Q8ICP9	Q8ICP9 plasmodium
20	189.5	5.0	2771	5 Q26216	Q26216 plasmodium
21	189	5.0	1936	5 Q815A6	Q815A6 plasmodium
22	189	5.0	2792	5 Q814R2	Q814R2 plasmodium
23	188	5.0	9904	16 Q8NWQ6	Q8NWQ6 staphylococ
24	187.5	5.0	1681	5 Q81548	Q81548 plasmodium
25	187	5.0	2747	5 Q8BJX9	Q8BJX9 plasmodium
26	186.5	4.9	1301	5 Q8IHQ2	Q8IHQ2 plasmodium
27	185.5	4.9	2235	5 Q81B09	Q81B09 plasmodium
28	185	4.9	2259	5 Q26223	Q26223 plasmodium
29	184.5	4.9	2529	16 Q25579	Q25579 helicobacte
30	184	4.9	1859	5 Q8IC27	Q8IC27 plasmodium
31	184	4.9	3504	5 Q81L45	Q81L45 plasmodium
32	183.5	4.9	1125	16 Q98PM9	Q98PM9 mycoplasma
33	183	4.9	3203	5 Q7YWE7	Q7YWE7 plasmodium
34	183	4.9	3203	5 Q7YWE6	Q7YWE6 plasmodium
35	183	4.9	4152	2 Q9ZHL3	Q9ZHL3 haemophilus
36	183	4.9	4152	16 Q7VLE8	Q7VLE8 haemophilus
37	182.5	4.8	2444	16 Q81GX1	Q81GX1 bacillus ce
38	182	4.8	6713	16 Q99U54	Q99U54 staphylococ
39	182	4.8	6713	16 Q931R6	Q931R6 staphylococ
40	181.5	4.8	4095	2 Q7WZ13	Q7WZ13 haemophilus
41	181	4.8	2563	5 Q8IHV4	Q8IHV4 plasmodium
42	181	4.8	2647	5 Q9U4X0	Q9U4X0 plasmodium
43	181	4.8	10061	5 Q813Z1	Q813Z1 plasmodium
44	180.5	4.8	149	2 Q9RM77	Q9RM77 clostridium
45	180.5	4.8	1113	5 Q815Q0	Q815Q0 plasmodium

ALIGNMENTS

RESULT 1

Q937W2	PRELIMINARY;	PRT;	741 AA.
ID	Q937W2		
AC	Q937W2;		
DT	01-DEC-2001 (TReMBLrel. 19, Created)		
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)		
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)		
DE	Fag protein (fragment).		
GN	PAG.		
OS	Bacillus anthracis.		
OG	Plasmid pXO1.		
OC	Bacteriia; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1392;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Ferrara;		
RA	Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,		
RA	Fasanella A., Francia M., Ciuchini F.,		
RT	"Sequence analysis of the genes encoding for the major virulence		
RT	factors of bacillus anthracis vaccine strain 'Carbosap'";		
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBSJ databases.		
DR	EMBL; AJ413937; CAC93935.1; -		
DR	GO; GO:0005576; C:extracellular; IEA.		
DR	GO; GO:0046821; C:extrachromosomal DNA; IEA.		
DR	GO; GO:0015070; F:toxin activity; IEA.		
DR	GO; GO:0009405; P:pathogenesis; IEA.		
DR	InterPro; IPR003896; Anthrax toxinB.		
DR	Pfam; PF03495; Binary_coxB; I.		
DR	PRINTS; PR01391; BINARYTOXINB.		
FT	NON_TER	1	741
FT	NON_TER	741	741
SQ	SEQUENCE	741 AA;	83153 MW; C7F95820B73065C0 CRC64;

Query Match 97.9%; Score 3695; DB 2; Length 741;
Best Local Similarity 99.6%; Pred. No. 7.4e-168;
Matches 719; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLLYFSDLNFOAPMVVTSITGDLSPSSELENIPSENOYF 60

Db 20 EVKQENLLNESSESSQGLGYFSDLNFAQPMVVTSSITGDLSPSELENIPSENOYF 79
QY 61 QSAIWSGFIKVKKSDDEYTFATSNADNHVTWDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 80 QSAIWSGFIKVKKSDDEYTFATSNADNHVTWDDQEVINKASNSKIRLEKGLYQIKIQY 139
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTSGPTVPDRDN 180
Db 140 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTSGPTVPDRDN 199
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYDFEKT 240
Db 200 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYDFEKT 259
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 260 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 319
QY 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 320 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 379
QY 361 NANIRYNTGTAPIYVNLPTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 380 NANIRYNTGTAPIYVNLPTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 439
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGNSWSEV 480
Db 440 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGNSWSEV 499
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
Db 500 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 559
QY 541 QYQKDIPTDFNFDQQTSONIKQLAELNATNIYTVLDKIKLNKWNILIRDKRPHYDR 600
Db 560 QYQKDIPTDFNFDQQTSONIKQLAELNATNIYTVLDKIKLNKWNILIRDKRPHYDR 619
QY 601 NNIIVAGDESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 620 NNIIVAGDESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 679
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 680 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 739
QY 721 IK 722
Db 740 IK 741

RESULT 2
Q937W3 PRELIMINARY; PRT; 743 AA.
AC Q937W3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE pag protein (Fragment).
GN PAC
OS Bacillus anthracis.
OG Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Carbosap.
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
RA Fasanella A., Francia M., Ciuchini F.;
RA "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus anthracis vaccine strain 'Carbosap'";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ413936; CAC93934.1; -.
DR GO: GO:000576; C:extracellular; IEA.
DR GO: GO:0046821; C:extrachromosomal DNA; IEA.
DR GO: GO:0015070; F:toxin activity; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR003896; Anthrax toxinB.
DR Pfam: PF03495; Binary toxB; I.
DR PRINTS: PR01391; BINARYTOXINB.
KW Plasmid.
FT NON_TER 1 1
FT NON_TER 743 743
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5DESA6 CRC64;
Query Match. 97.9%; Score 3695; DB 2; Length 743;
Best Local Similarity 99.6%; Pred. No. 7.4e-168;
Matches 719; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENLLNESSESSQGLGYFSDLNFAQPMVVTSSITGDLSPSELENIPSENOYF 60
Db 22 EVKQENLLNESSESSQGLGYFSDLNFAQPMVVTSSITGDLSPSELENIPSENOYF 81
QY 61 QSAIWSGFIKVKKSDDEYTFATSNADNHVTWDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 82 QSAIWSGFIKVKKSDDEYTFATSNADNHVTWDDQEVINKASNSKIRLEKGLYQIKIQY 141
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTSGPTVPDRDN 180
Db 142 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTSGPTVPDRDN 201
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYDFEKT 240
Db 202 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYDFEKT 261
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 262 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 321
QY 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 322 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 381
QY 361 NANIRYNTGTAPIYVNLPTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 382 NANIRYNTGTAPIYVNLPTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 441
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGNSWSEV 480
Db 442 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGNSWSEV 501
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540
Db 502 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 561
QY 541 QYQKDIPTDFNFDQQTSONIKQLAELNATNIYTVLDKIKLNKWNILIRDKRPHYDR 600
Db 562 QYQKDIPTDFNFDQQTSONIKQLAELNATNIYTVLDKIKLNKWNILIRDKRPHYDR 621
QY 601 NNIIVAGDESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 622 NNIIVAGDESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 681
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 682 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 741
QY 721 IK 722
Db 742 IK 743
RESULT 3
Q32739
ID Q32739 PRELIMINARY; PRT; 876 AA.

QY	573	NIYTVLDTKILNAKNILIRDKRFHY---DRNIIAAGADESVVKEAHREVINSSTGL---	627
Db	599	KIYNV---KLERGMNILLKTPTFTFDYNNYP--STWSNVNTTNOGLQGSANKJNG	652
QY	628	-----LLNIDKURKILSGV-----IVEI-----EDTEGL-----KEVI	656
Db	653	ETWKIPMBELPKYRYVSGYSGKDLPTNSIIVKIAKEEKTDILVPQGYTKFSYEPE	712
QY	657	NDRYMLNIS-SLRQDGKTFIDFKKYN--KULPYISNPNYKV-----NVYAVTKENT	706
Db	713	TTEKDSNTEITLIGSGTYVLDNLSTELNSTPEILDEPEVKIPTDOEIMDAHKIYFADL	772
QY	707	IINPSENGDTSTNGI	721
Db	773	NFNPS-TGNTYINGM	786
RESULT 4			
ID	006498	PRELIMINARY; PRT; 879 AA.	
AC	006498;		
DT	01-JUL-1997 (TrEMBLrel. 04, Created)		
DT	01-JUN-1998 (TrEMBLrel. 05, Last sequence update)		
DE	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Sb component.		
GN	SBS.		
OS	Clostridium spiroforme.		
OC	Bacteria; Firmicutes; Molllicutes.		
NCBI_TaxID=29348;			
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=CS246;		
RA	Gibert M., Perelle S., Daube G., Popoff M.R.;		
RT	"Clostridium spiroforme toxin genes are related to C. Perfringens iota		
RT	toxin genes but have a different genomic localization.";		
RL	Syst. Appl. Microbiol. 20:337-347(1997).		
DR	EMBL; X97969; CAA66612.1; --		
DR	HSSP; P13423; 1ACC.		
DR	GO; GO:0005576; C:extracellular; IEA.		
DR	GO; GO:0015070; F:toxin activity; IEA.		
DR	GO; GO:0009405; P:pathogenesis; IEA.		
DR	InterPro; IPR003896; Anthrax_toxinB.		
DR	Pfam; PF03495; Binary_toxB; I.		
DR	PRINTS; PR01391; BINARYTOXINB.		
QY	SEQUENCE 879 AA; 98739 MW; 40685ACB8E05BA01 CRC64;		
Query Match			
Best Local Similarity 28.8%; Pred. No. 4.4e-34;			
Matches 260; Conservative 147; Mismatches 258; Indels 237; Gaps 34;			
QY	4	QENRLNSESSSQGLGYFSDLNFPQPMVVTSTTGDISIPSSSELENIPSENQ-YFQS	62
Db	36	QEVETTNKTVSSNGLMGYFADFHFKDLLEMAPVKNGLKPEKNKVEKLLTEKTNKS	95
QY	63	AIVSGTIKVKSDBYFPAISADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIQYOR	122
Db	96	IRVTGRIIPSGDGYTLSTDKDN-VLMQINAE---GEIANTLKNMTKQGEYSIRIEIQ-	150
QY	123	ENPTREGKLDP-----KLYWTDSONKKEVISSDNLQLPELKQSSN-----	162
Db	151	-----DKDIGVDDLSSEKLYW-ELNGDKTLIPKULFLRYSKIDENDPFPKDNFFDLK	205
QY	163	--SRKKSTGA-QPTVPDRNDGIPDSLEVEGYTVVKNRTFLSPWISNIHEKGLTKY	219
Db	206	LKRSARLASGWDGDELDTDNNDIPAYEKNGYTI-----KDSIAVKW-EDSFAQQGYKY	260
QY	220	KSSPEKWTASDPVSDPEKVTGRIDKNVSPEARHPLVAAPIVHVDVMEIILSKNEQST	279
Db	261	LSSVLENTAGDPYTDYQKASGSFDKAKAEARDPLVAAPVVGVGHEKLIISTNEHAST	320
QY	280	QNTDSETRTISKNTSTRT--HTSEVHGNAEVHASFPDIIQGSVSAGFSNNSSTVAIDHS	337

```

Db 321 ----DQKTVSRNTNSKTDANTAGVAINAYQNGF-----TGSITTVNYSHTTENSTAVQNS 373
Qy 338 LSLAGERTWAETMGLNTADTARLANANIRYVNTGTAPIYVNLPTTSLVGLKNGQTLTIKAK 397
Db 374 ---NGE-SWNTSLINKGESAYINANVRVYNTGTAPYKVTPTTNLVL-DGDTLTITIKAQ 428
Qy 398 ENQLSQILAPNNYPSKNLAPALNAQKDFSTPTITVNNVQFLELEKTKQLRLDQDOVYG 457
Db 429 DNOIGNNLSPNETYPRKGLSPALNTMQFSLRIPDINVDQKLDAGKQIKLETTQVSG 488
Qy 458 NIATYNFENGVRVDT-GSNMSEVLPIQIETTARIIFN-GKDLNLVERRIAAVNPSPLE 515
Db 489 N---VGIRKSQOQIITEGNSWSDYISQISLSASIILDTGSD--VFERRVTAKDSSNPED 543
Qy 516 TTKPDMTLKALKIAFGNEPNNGNLQYQKQITE--FDNFEDQOTSONIKQLAELNATN 573
Db 544 KT-PVLTIGEALKEKAPGATNGEILFYNGMPIDESCVELIFDGTANLIKRLNALNDKK 602
Qy 574 IYTVLDKIKLNAMNILLRD-----IAVGADESVMKAEHREV-----KRFHY----- 598
Db 603 IYV-----QLERGKMLIKTSTFYFNNFDGYNVFPSSWSDVSNQDGLQNAANKLSGETK 658
Qy 594 -----KRFHY----- 598
Db 659 IVPMSKLNPKRYVPFSGYLKNSSTNPITVNIKAKQKTYNLVSENDYKFSYEFETIG 718
Qy 599 -DRNN-----IAVGADESVMKAEHREV-----I 620
Db 719 RASNEIETLTSSGTFIDNLNLSITELNSTPEILKSPDIKVPSDQELI-DAHKYVADLSF 777
Qy 621 NSTEGELLN-----IDKDIRKILSGYIVIEB-DTEGLKEVINDRYDMLNINSLRQD 671
Db 778 NOSTANYLDGLYFEPTQNKVELDYIQKYKVEATLEYSGFKDICTGDKELRNVTGDSNQ 837
Qy 672 GKT-FIDKPKYNDKPLXYISPNY---KVVYVAVTKENTIINPSENGDTSNGFKKILI 726
Db 838 PKTYNVNFRSY-----FISGENVMPYKLRIVATPEN-----KELLV 875
Qy 727 FS 728
Db 876 LS 877

RESULT 5
ID Q9KH41 PRELIMINARY; PRT; 876 AA.
AC Q9KH41;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CdtB.
GN CdtB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 20309;
RA Chang S.Y., Song K.P.;
RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain
RL CCUG 20309."
RM Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF271719; AAF81761.1; -.
DR HSPSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB; I.
DR PRINTS; PR01391; BINARYTOXINB.
DR SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;

```

```

Query Match 23.4%; Score 881.5; DB 2; Length 876;
Best Local Similarity 31.6%; Pred. No. 5.5e-34;
Matches 251; Conservative 140; Mismatches 279; Indels 125; Gaps 30;
Qy 4 QENRLNESSESSOGLLYGYSFDLNFQAPMVVTSSTGDLSPSSSELEN-IPSENYQFOS 62
Db 40 KKKEIVNEIDILPNGLMGYYFTDEHFQDLKMAPKIDGNLKEPEEKVKLLDKDSDVKS 99
Qy 63 AHWGFIKVKSDSEYTFATSAADNHVTMVDDQEVINKASNSNKHIRLEKGLYQIKIYQR 122
Db 100 IRWTGRIIISKOGGEYTLSTRDD-VLMQVNTSTI---SNTLKVNMKKGEYKVLIELQD 155
Qy 123 EN--PTEKGLDFKLYWTDSONKEVISSDNQLPELKQKSSNRKRSTSAQTPV---- 176
Db 156 KNLGSIIDNLSPNLY-ELDGMKKIPEENLFLRDY-----SNIKO-----DPFIPNNNF 205
Qy 177 -----DADNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSS 222
Db 206 FDFKMSDWEDELDYDNDNIPDSYERNGYTI-----KOLIAVKWEDSFAE-QYKKTIVSN 260
Qy 223 PEKWSSTASDPYSDFEKVGTGRIDKQVSPPEARHPLVAAYPVHVDMENIILSKNEDOSTONT 282
Db 261 YLESNTAGDPYTDYERKASGDFKAIKTEARDPLVAAYPVGVGMEKLIISTNEHAST--- 317
Qy 283 DSERTISKNTSTSRTHTEVHGNAEVHASPFDIGGSVAGFSNSNSTVAIDHS----- 337
Db 318 -DQKTVSRATNKSNTA-----GVSNVGYQNGFTANVTNYSHTTDN 363
Qy 338 ---LSLAGERTWAETMGLNTADTARLANANIRYVNTGTAPIYVNLPTTSLVGLKNGQTLTI 394
Db 364 STAVQDSNGESWNTGSIKNGESAYINANVRVYNTGTAPYKVTPTTNLVL-DGDTLSTI 422
Qy 395 KAKENQLSQILAPNNYPSKNLAPALNAQKDFSTPTITVNNVQFLELEKTKQLRLDQO 454
Db 423 KAQENQIGNNLSGDTYPPKGLSPLALNTWDQFSSRLIPINDYDQKLDAGKQIKLETTQ 482
Qy 455 VYGNIAVTYPPENGRVVDVGTGSNWSEVLPOIQTARIIFNGKDLNLVERRIAAVNPSPDPL 514
Db 483 VSGNFGTKN-SSQGI-VTEGNSWSDYISQISASIIIDTEN-ESYERRVTAKNLQDPE 539
Qy 515 ETTKPDMTLKEALKIAPGFPNEPNNGNLQYQKQITE--FDNFEDQOTSONIKQLAELNAT 572
Db 540 DKT-PELTIGEALKEKAPGATKKDGLLYFNDIPIDESCVELIFDDNTANKIKDLSKLTSDK 598
Qy 573 NIYTVLDKIKLNAMNILLRDKEFHY---DRNNIAGADESVVKEAHRVINSSTEGE-- 627
Db 599 KIYV-----KLERGMLIKTPTTYTTFNFDYNNYP--STWSNVNTTKDGLQGSANKLNG 652
Qy 628 -----LLNIDKDIRKILSGY-----IVEI-----EDTEGL-----KEVI 656
Db 653 ETKIKIPMSSELKPKYKRVFSGYSKDPKLTNSIIVKIKAKEEKTDLVPEQGYTKFSYEFE 712
Qy 657 NDRYDMLNIS-SLRQDKTFIDFKYND--KLPLYISNPNYK-----VNVYAVTKENT 706
Db 713 TTEKSSNIETILIGSGTIVLDNLSTELNSTPEILDEPEVKIPTQOBIIDAKHIYFADL 772
Qy 707 IINPSENGDTSNGI 721
Db 773 NFNPS-TGNTYINGM 786

RESULT 6
ID Q46221 PRELIMINARY; PRT; 875 AA.
AC Q46221;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Iota toxin component Ib precursor.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;

```

RN	SEQUENCE FROM N.A.	586	IIKEQKYLDDKKIYV	-----	KLKSGMILLIKVPSYFTNFBYNNFP	-----	ASWSNIDTKNQ	639		
RP	STRAIN=NCIB 10748;	618	EVINGSTGL	-----	LLNIDKDIRKILSGY	-----	IVEIEDTGLKEVIND	658		
RX	MEDLINE=94041637; PubMed=8225592;	640	DGLQSVANKLSGETKIIIPMSKLPKRYVFGSKDPSKTSNITVNIKSKQKTDYLP	-----				699		
RA	Perelle S., Gibert M., Boquet P., Popoff M.R.;	659	RYDLNIS	-----	SLRQDKTIDPKYND	-----	KLPIYISNPYKV	696		
RT	"Characterization of Clostridium perfringens iota toxin genes and expression in Escherichia coli.";	700	EKDYTKFSYEFETGKSSDIITLTSSGVIFLDNLSITELNSTPEILKEPIKVPDQE	-----				759		
RI	Infect. Immun. 61:5147-5156(1993).									
RN	SEQUENCE FROM N.A.	697	-----	NVAVTKENTIINPSENGDTSTNGI	721					
RP	STRAIN=NCIB 10748;	760	ILDAHKKYADIKDI	-----	NTGNTYIDGI	785				
RA	Popoff M.R.;									
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBSJ databases.									
DR	EMBL; X73562; CAA51960.1; -									
DR	PIR; I40862; I40862.									
DR	HSP; P13423; 1ACC.									
DR	GO; GO:0005576; C:extracellular; IEA.									
DR	GO; GO:0015070; F:toxin activity; IEA.									
DR	GO; GO:0009405; P:pathogenesis; IEA.									
DR	InterPro; IPR003896; Anthrax toxinB.									
DR	Pfam; PF03495; Binary toxinB; I.									
DR	PRINTS; PR01391; BINARYTOXINB.									
KW	Signal.	34	38	POTENTIAL.						
FT	SIGNAL	212	875	IOTA TOXIN COMPONENT IB.						
FT	CHAIN	212	875	IOTA TOXIN COMPONENT IB.						
SQ	SEQUENCE	875	AA; 98468	MM; C9AE092CD3818921	CRC64;					
	Query Match	23.2%	Score 877;	DB 2;	Length 875;					
	Best Local Similarity	31.2%;	Pred. No. 9e-34;							
	Matches	253;	Conservative	131;	Mismatches	277;	Indels	150;	Gaps	31;
QY	1	EVKQENRLNESSSQGLLYGYSFNLNFCAPMVVTSSTGDLSPSSSELENIPSE-NOY	59							
DB	35	DTNQKEEITNENTLSSNGLMGYFADEHFKDLNAPIKNGDLKFEKVKVLDLTDSS	94							
QY	60	FQSAIWGFTKVKKSDYVFAFSADNHVTWMDQEVINKASNNKIRLEKGLYQIKI	119							
DB	95	IKSIRWTGRILPSEDGEYILSTR-NDVLMOINAKGDIK- --TLKVNKKGCAYNRIE	150							
QY	120	YQREN-----PTEKGLDFLYWDSQNKKEVSSDNQLPELKQKSSNRKST	170							
DB	151	IQDNLGSIDNLSP-----KLYW-ELNGNKTVIPENLFFRDYSKIDEND-----	195							
QY	171	AGTVP-----DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIH	211							
DB	196	--PFIPIPNFFDYRFFSAWEDDLTDNDNIPDAYEKNGYTI- --KDSIAVKNDSFA	249							
QY	212	EKKGLTKYSSPEKWTASDPYDFKVTGRIDKNVSPARHPLVAAYPIVHVDMENIL	271							
DB	250	E-QYKKYVSSYLESTAGDPYDYQKAGSIDKAIKLEARDPLVAAYPVVGVGMENLI	308							
QY	272	SKNEDQSTQNTDSTRITSKNTSTSRTHTSVHGNVHASFDPDGGSVSAGFSNSST	331							
DB	309	STNEHSS-----DOGKIVSRATTSKTDANTV-----GVSIAGYQNGFTGN	351							
QY	332	VAIDHS-----LSIAGERTWAEITGLTATARNANIRVYNTGTAPIVNLPTSL	383							
DB	352	ITTSYSHTDNSTATVQDSNGESNWTGLSINKGESAVINANRYNTGTAPYKVTPTNL	411							
QY	384	VLKXQTLATIKAKENQLSOILARNVYPSKNLAPALNAQDPSPTIMVYQFLE	443							
DB	412	VL-DGETLATIQAQDNIGNLSNETYPRKGLSPALNTMOQFNARLPINVDQKKLD	470							
QY	444	KTQLRLTDQVGNIAFYNFENGRVRVDTGSNMSEVLPOIETTARIIFNGKDLNVER	503							
DB	471	SGKIKLETTQVSGNYGTGN-SQOI-ITEGNSNSNVIQSDVSASIIID-TGSQTFER	527							
QY	504	RIAAVPSDPLETTKPDMLKALKIAPGNEPNGNLOY-QGKDITE--PDFNPDQTS	560							
DB	528	RVAKEQGNPDKT-PEITIGEAIIKAFSATK-NGELLYFNGIPIDESCVELIFDNTSE	585							
QY	561	NIKQLAELNATNTYVLDKIKLNAKXNIIIRKRF--HYDR-NNIAVGADESIVKEAHR	617							

```

QY 226 NSTASDPYSDPEKVTGRIDKQVSPPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSE 285
DB 221 PCTANDPYTDEKVSQGLDPSVSMVARDPMISAYPIVGQMERLVVSKSE---TTIGDS 286
QY 286 TRTISKNTSTRTHTSEYHGAHVHAFFFDGG-----SVSAGFSNNSSTVAIDHSL 338
DB 287 TKSMKSTSHSSTNTV--CAEVSGLQLAGGIFPVFSASANTSHTWNTSTVDDT- 343
QY 339 SLABERTWAEFMGLTADTARLANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKE 398
DB 344 --TBE-SFSQSLSTNGESAVINIRYVNTGTAPIYVNLPTTIIVIDK-QSVATIKGE 399
QY 399 NOLSOILAPNNYPSKNLAPIALNAQKDFSTPTIMNYNOFLEKTKQLRLDQVGN 458
DB 400 SLIGDYNFGGTYPIEIGPPMALNTMDOFSRLIPINYNQKSIDNGGTVMLSTSOFTCN 459
QY 459 LATYNFENGRVVDGTGNSWSELVQIOBTARII--FNGKDLNVERRAIAAVNPSPLET 516
DB 460 FAKYN-SGNLUTD-GNNWGYLGTIKSTIASLTLSFGGTTQVA---VVPNFSPEBK 514
QY 517 TKPOMTLKEALKIAPGNEPNNGNIQYQKQIT---EFDFNPDQOTSONIKQNAELNATN 573
DB 515 T-PKLTLEQALVKAFALFKKNGKGFYFHGLEISKNEKIQVFLDNTNNDFENQLKNTADKD 573
QY 574 IYVLDKIKLNAKNMILIRDKRHHYDRNIAV 605
DB 574 IMHCI--IKRN--MNILVKVITPKENISSINI 601

RESULT 8
Q844J8 PRELIMINARY; PRT; 775 AA.
AC Q844J8
DT 01-JUN-2003 (TremBrel. 24, Created)
DT 01-JUN-2003 (TremBrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
DE VipiAC.
GN VipiAC.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y., Chen J., Pang Y.;
RT "Cloning of vip1A(C) and vip2A(C) from Bacillus thuringiensis HD201.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DSJ databases.
DR EMBL; AY245547; AAO86514.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anchrax_toxinB.
DR Pfam; PF03495; Binary_tox3; I.
DR PRINTS; PRO1391; BINARYTOXINB.
DR SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;

Query Match 19.5%; Score 734; DB 2; Length 775;
Best Local Similarity 29.2%; Pred. No. 4.8e-27;
Matches 225; Conservative 133; Mismatches 292; Indels 120; Gaps 30;

QY 4 QNRLINSESSQGLGYFSDLNLFQAPMVVTSSTGDLISPSSELEN--IPSENYQYQ 61
DB 42 QKNQ---QKEMDRKGLGYFGKDP-SNLTMFAPTRDNTLIYDQQTANKLLDKQQQYQ 97
QY 62 SAWSGFIKVKSDYTTATGADNHVTWVDQSVINKASNSKIRLEKGRLYQIKIYQ 121
DB 98 SIRWIGLIQSKKEGDFTNLSEDEQAIIIDGKIISNKGKQVHVLEKVLPIKIEYQ 157
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISSNLQLPKLKQSSN----- 162
DB 158 SD--TKFNIDSKTFKFKFLKFDKSDQSQVQRDELNPFNFKESSEFLAKASKTNFFM 215
QY 163 SRKKGSTAGPVPDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSS 222

```

```

DB 216 QKMKRDIDED---TDTGDSIPDLMEENGYT--IQNK--VAVKM--DDKPAQOQYKYLSS 267
QY 223 PEKWTASDPYSDPEKVTGRIDKQVSPPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNT 282
DB 268 PYQAHVTGDDYTDWEKAAGDIPKSNAAATRNPLVAAPFSINVDKMKILSKDSNLS---- 323
QY 283 DSETRTISKNTSTRTHTSEYHGAHVHAFFFDGGISVSAGFSNNS--TVAIDHSL 339
DB 324 -----NSAAHNSNYTYANSE-----GASIEAGFGPKGFGVGSANTYQHTET 366
QY 340 LAGERTW---AETMGLNTADTARLANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIK 395
DB 367 VGSQ--WGNKSKSTEQFNSASAGYLNANVHYNNVGTGGIYDAQPTTSFIL-QDSTIATIT 423
QY 396 AKENQLOILAPNNYPSKNLAPIALNAQKDFSTPTIMNYNOFLEKTKQLRLDQV 455
DB 424 AKSNATALSIPSGDRYPASK-BGISLKTWDDFNHPITLTKPQLDAVLNNEVIKINTDOT 482
QY 456 YGNATYNFENGRVVDGTGNSWSELVQIOETARIIFNGKDLNVERRAIAAVNPSPDPLE 515
DB 483 DGRYGIIGVDG---KAEIGDRWSPIIDEIKGTASIIIDPADGKALETIRIAAKDYKNPBD 539
QY 516 TTKPDMTLKEALKIAPG--FNEPNGLQYQKD---ITEFDN-----FDQOTSON 561
DB 540 KT-PSLTIKELGLIAYPESISEDKGILFYEYKNDGKVTKKQLSEENIMPYLDEDTSK 598
QY 562 IKQNAELNATNIYTVLDKILNAKNMILIRDKRHHYDRNIAVAGADESVVKEAHEVIN 621
DB 599 FERQLSDGSAGLY---DIKLTPKKNITIR-----LATVTLGFDQFQSAYPWENATW 647
QY 622 SSTEGIL-----INIDKDIRK-----ILSGYIVE--IEDTEGLKEVI---NDR 659
DB 648 SDRFNLGLSLAIPOESKYTTPKVKKNYDYLTIGYIKHDFTTDNESLGIIVATFKKN 707
QY 660 YDMLNITSSLRQDGKTFIDFKYNDKULPLYISNPKNYKNVYATKENTIN 709
DB 708 FEMNMGTSIPSONSGEFGKFTIKTQ-NISG-DYILDSIQLMKRNDVN 755

RESULT 9
Q8KYK2 PRELIMINARY; PRT; 225 AA.
AC Q8KYK2
DT 01-OCT-2002 (TremBrel. 22, Created)
DT 01-OCT-2002 (TremBrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBrel. 24, Last annotation update)
DE Protective antigen-related protein, (pXOI-111).
GN BXA0163.
OS Bacillus anthracis.
OC Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2012;
RC MEDLINE=22061436; PubMed=12004073;
RA Read T.D., Salzberg S.L., Pop M., Shumway M., Unayam L., Jiang L.,
RA Holtzapple E., Bush J.D., Smith K.L., Schupp J.M., Solomon D.,
RA Keim P., Fraser C.M.;
RT "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in
RT Bacillus anthracis."
RL Science 296:2028-2033(2002).
DR EMBL; AE011190; AAM26108.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 225 AA; 25402 MW; 2E121BES4295P9C8 CRC64;

Query Match 7.9%; Score 296.5; DB 2; Length 225;
Best Local Similarity 37.2%; Pred. No. 6.1e-07;
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;

QY 566 LAELNATNIYTVLDKILNAKNMILIRDKRHHYDRNIAVAGADESVVKEAHEVINSSITE 625

```

Db 1 MESLGINNINVALDRIKNAKNNILVRDP-YHYDNGNNGVDDSYLKNAKQILNWSDD 59
Qy 626 GLLNIDKIDKILSGYIVIEDTE-----GLKEVINDRYDMLNSSLRQDG 672
Db 60 GVSLLNEDVNAQSGYMLQIKFSPNHLTNSPVITITLAGKDSGVGELYRVLS-----DG 113
Qy 673 KTFIDFKYNDKPLIYINPNKVNVAATKEN-TIINPSENGDTSTNGIKKILIFSKKG 731
Db 114 TGFIDFKFENWRSVLV-DPGDDVYVAVTKEDFNATRDENGIA-NKLKNTLVLSGKI 171
Qy 732 YEI 734
Db 172 KEI 174
RESULT 10
Q8RGK2 PRELIMINARY; PRT; 1881 AA.
AC Q8RGK2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hemolysin.
GN FN0291.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ARCC 25586;
RX MEDLINE=2186394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fomstein M., Kyrpides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
J. Bacteriol. 184:2005-2018(2002).
RL EMBL: AEO10541; AAL94497.1;
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR008638; Haemagg_act.
DR InterPro: IPR000508; Peptidase_S26.
DR InterPro: IPR001452; SH3.
DR Pfam: PF05860; Haemagg_act; 1.
DR Pfam: PF00018; SH3; 1.
DR PROSITE: PS00761; SPASE_1_3; 1.
KW Complete proteome.
SQ SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;
Query Match 5.8%; Score 220.5; DB 16; Length 1881;
Best Local Similarity 20.8%; Pred. No. 0.04;
Matches 175; Conservative 131; Mismatches 303; Indels 233; Gaps 37;
Qy 30 QAPWVTS-----STTG-----DLSIPSELE---NIPSENY-FQSAIWSGFI 69
Db 246 QAPYATKGVGVSSKGVYKDTQAKRDKISSTETETIGSKLLAENAIKSGKTSNG 305
Qy 70 KVKSDVETATSDADNHTWVDDQEVIN-----KASNSKIRLEKGR 112
Db 306 QIRANNITINGVDSNLIPTNKDITISGNLKNMGVSSNINLVKEIENSKNVWVEE-K 364
Qy 113 LYQKI-----QYQENPTKGLDFKLYWT--DSQNKKEVSSDNLQLPELKSS 161
Db 365 LSSTKITNLGNLSAKEIEKTIFNSGKLFSGKITAKDFKNGEV-SSENITTTNLENSK 423
Qy 162 -----NSRKRKTSAGTVPDRNDG----- 182
Db 424 INVKENINSINVKNTNAEITISKLNNTNLDNRGNITIIINNVSSTVANNKGLVGNNTIN 483

Qy 183 ---IPDSLEVEGYVDVVRKRTFFSPWISNTHHEKGLTKYKSSPEKWTASTDPYSDPEKV 239
Db 484 SQNLNTATVQKTLIDIKNKINSGKILSDNLTDXIFSSGNISAKVITTOELINSGETI 543
Qy 240 TGRIDKNVSPEARHPLVAAYPIVHVDMEN-----IILSKNEDQSTQNTDSTRITSKNT 293
Db 544 SNNLSNNINNNKNIIFVGNLKIISNNLNSGVIEGLELNTSIENTGNITIKNKLTSQNL 603
Qy 294 STSRTHTEVHGNAEVAHSAFFDGGVSAGFSNSSTVA---IDHSLSLAGERTWAETM 350
Db 604 NNKK-----NTANVAGFLDVHNKISS-VGNIKAITMKTNNLNSGNIL-----TN 648
Qy 351 GLNTAD-----TARLNANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQ 403
Db 649 SLTTAENINKSITAKNISQNLVNSGS-----VISDNIITVA-----DN 687
Qy 404 ILAPNNYPSKNIAPIALNAQKDFSTPTIMYNOFLELEKTKQLRLDQDQVYGNITATYN 463
Db 688 ITNTNIF-----ANEKISADKIS-NSNK-LVAKNTEITKLTND--GNIWVK- 730
Qy 464 FENGVRVDTGSNWSEV-----LPQIQTETARIIFNGKDLNLFVERIAAVNPSPDPLE 515
Db 731 -ENLKAKDITNSNTIKVGNLNTDKLQNSKTLIAKNINIEKSLNNGINKITSLNANINTS 789
Qy 516 TTRPDMTLKEALKIAFGFNEPNGNLQYQKDKITEFDFNFDDQTSQNIKNQLAELNATNY 575
Db 790 DIKNNGIIOAIK-----NINIKTSNDLKLKGKYTANDSLNINAKSLENNGN--- 836
Qy 576 TVLD---KIKLNAKQILIRDKFHDHNNIAV-----GADESIVYKEAHREVINSS--T 624
Db 837 --LENDGKIKFNLTKNLNNKISSSSNLNITANEISNNGNVNSIIIGSEANLIITANSLKN 894
Qy 625 EGUILL-----NI-DKDIRKILSGYIETEDTEGLKEVINDRYDMLNSSLRQ 670
Db 895 EGNLLFQEGENLKYTKTNTGTNTGVSSLGKLEAKDVNDKHIISDNDLTIDVNSITN 954
Qy 671 DGKTF-----IDPKK--YNDKLPYISNPYKNYVAVTKENTIIINPSENGDTSTNGIK 722
Db 955 KGLLYSTNNMKVDFKFNFLDKAIIYSSG-----DITIN-SENG-TFTNRVG 999
Qy 723 KI 724
Db 1000 DI 1001
RESULT 11
Q8II104 PRELIMINARY; PRT; 3468 AA.
ID Q8II104
AC Q8II104;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pf11_0371.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.B., Bowman S., Paulsen I.T., James K.,
RA Bisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Sun B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.,
RT "Genome sequence of the human malaria parasite Plasmodium
falciparum."
RL Nature 419:498-511(2002).

DR EMBL; AB014841; AAN35955.1; --
KW Hypothetical protein.
SQ SEQUENCE 3468 AA; 411913 MW; 506F7D62999BA7B1 CRC64;

Query Match
Best Local Similarity 5.8%; Score 217.5; DB 5; Length 3468;
Matches 159; Conservative 131; Mismatches 274; Indels 271; Gaps 37;

QY 55 SENQYFQSAIWSGFIKVKKSDEYVTSADNHTVMWDDQEV---INKASNSNKIRLSEK 111
DB 160 SENEY-----TKNDKYNSTNSGY-----NKEIEFLKNAKESLSKKIEKN 202

QY 112 RLYQIK-----TOYQRENTPEKGLDPK---LYW---TDSQN-----KKEVTSNNL 151
DB 203 INLLKCDSDNNDLNLEFKKNVHLNNDYSEDLFDNRVNNQNNKLDIIKDYIHHNNI 262

QY 152 QLPKQKSGNSRKKRSTAGPTVPDRDNDGIPDSLEVEGYVDV-KNKRTFLSPWISNI 210
DB 263 MI-----EENNDNRNDKCNFKNTNPNVERHIIVDTKNGNANNINSNNFDS--ISNI 315

QY 211 HEKGLTKYKSSPEKSTASDP-----YSDFE-KVTGRIDKXVSPARHPLVAAYPIVHV- 264
DB 316 NKK-----ISYPINMYSNSEDKTLNQINLWDSILSSDSLKNAYSFYLS 360

QY 265 ---DMENIILSKNEDQSTQNTD--SETRTISKNTSRTHTSEVHGNAEVAH--SFFDIG 317
DB 361 EKNSNNGIKNSHKNNRYRVIDEDINISFESNISNLHDSMDNNEPNSLCSF---- 416

QY 318 GVSAGFSNSSTVAIDHSLSLAGERTWAETVGL-----NTADTARLANIRY---- 366
DB 417 -----NSSSKCVXDEMITQVGNRKVKYKVTMDANDNMKNSTEQHSINMESHFKNQ 468

QY 367 -VNTGTAPIVNVLPVTLVLGKNQ-----TLATIKAKE--NOLSQLAPNN--- 409
DB 469 KINSKEDANDNLSHILNNKGIGQVNSLSDYLSIKKQESNNNSNNEALNNIN 528

QY 410 -----YYPSSKNLAPI-----ALNAQKDFSTP--- 431
DB 529 TNNNGSSNNKNSNVYKTSQYVYNNENDINNHNHQLNLSYMKNSNNLNTSGFKKIPGN 588

QY 432 -----ITWNYQOFLEKTKQLRLTDQVYGNIAFYNFENGVRVDTGSNWSVLPQIQ 485
DB 589 KNIISIDFDNIFKSYIKENVVVKQESNINHOFVEKNTDEINIKN-----NNIENNTQ 644

QY 486 ETTAR-----IIFNGKDLNVERIAAVNSDPLETTK----- 518
DB 645 NITCNPNFINTDIIINKKIKIETKKIDSLILNDLSLKLBSLSSINIDRYTKNYEBK 704

QY 519 ---PDMTLKEAL-----KIAFGNEPNGNLQYQKGDITEFDFNFDQOTS 559
DB 705 FLDDVILDDSIATSNEHQSNYTTNTHIFDNNNNNNNNINQEKDLFQNDYN---KET 761

QY 560 QNIKNQALAEALNATNIYVLDKIKLNAMNIIIRDKRPYDNRNIAVGADESVVYKAEHREV 619
DB 762 YNNNIMSENNAENLF-----KISYSCNDLVLGKNEIILD--RNV 800

QY 620 INSGTEGLLNDIKIRKILSGVIEIETDEGLKE-----VINRDYDMLNIS--L 668
DB 801 ENSKTE-----QVEEYQNKEDIKLYHKDDNIIIDNNYNDVNNIKKDCHL 844

QY 669 RQDKTEFIDFKYND--KLPL-----YISNPNYKNV--VYAVTKENTIIINPSEN 713
DB 845 KMDNQDNINKSQDSKKNKPINNECNVIHKNKFNKINEIDHFKENTILNYSKN 899

RESULT 12
Q9PQ08
ID Q9PQ08 PRELIMINARY; PRT; 4688 AA.
AC Q9PQ08;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DR 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein UU482.

GN UU482.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serovar 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
urealyticum";
RL Nature 407:757-762(2000).
DR EMBL; AE002145; AAF30894.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 4688 AA; 534880 MW; B53ABFAFEE1997E CRC64;

Query Match
Best Local Similarity 5.4%; Score 205; DB 16; Length 4688;
Matches 174; Conservative 132; Mismatches 324; Indels 186; Gaps 40;

QY 19 LLGYFSDLPNQAPMVVSTTQGLIPSELENIPSENOYFQSAIWSGFIKVKKDEYV 78
DB 3699 LVDVYILD-NIHQNIIDETRIKFDHNV-SKEIENFGVTMISKHGWNKSPDTDTANFEFK 3756

QY 79 FATSADNHVTMWDDQEVINKASNSNKIRLEKGRLYQIKIQORENPTEKGLDFLYWTD 138
DB 3757 IETQ-----DDNDVLNNIDATVKFDEHNNIKQIVRIKEN-----ND 3795

QY 139 SQNKKEVSSDNLQPLPELKOKSSN-----SRKKSTAGTVPDRD----- 179
DB 3796 WLKGGQI---DNLN-PETKYKLENIELSKPLKTHLSVSINDKENISLITETGNPVLKV 3851

QY 180 ---NDGIPDSLEVEGYVDVQKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYS 235
DB 3852 IQQNDTINTQTTINTVLSGVNSK-YNGRQIKVYKDNVNNVYESS---LITLQKQND 3907

QY 236 FEKVTGRIDKXVSPARHPLVAAYPIVHVDMENIILSKN-EDOSTQNTQSET-RTISKNT 293
DB 3908 YQLLSNLN-----REYRPEKIEIHISNTNPNFEDLEKLGVSNTFITQKNT 3957

QY 294 STSEHTS-EVHGNAEVAHSP-----FDIGGSVSAGPS-----SNS 335
DB 3958 TVQWNSSATIVTGRVGNFNFKISBEDKILENQVQVWVAFKETTIRDTNTWQYTRPLK 4017

QY 336 HSLSLAGERTWAETMGLNT---ADTARLANIRYVNTGTAPIVNVLPVTLVSLVKGKQTLA 392
DB 4018 DVTDFKEGTWAHDLNSNVNFKETTYKLVKIQVKNKTKAKNNNSNNVILDTNISI 4077

QY 393 -----TIKAKENOLSOILAPNNYPSKNLAPALN-AQKDFSTPTWNY--NQFLELE 443
DB 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQTINFVLSGVKSSWGVKKIKLSYKNDTSESI 4137

QY 444 KTKQLRLTDQVYGNIAFYNFENG-----VRVDTGSNWSVLPQIQETTARIIFNKGD 497
DB 4138 HTNEVLIESNKTQYNILNKLKRNRTYTLIDVKLIDNNVSDFPKGNLTNSFITRTSA 4197

QY 498 LNLVERRIAAVNSDPLETT-----KPDWTLKEA-----LKIAFGNEPNGNLQ 541
DB 4198 INVLNIEISNASTNLKSTIIKINLNDPDLVLRDKQDQATVYGNKQAMGFITVSGNIK 4257

QY 542 YQKGDITEFDFNFDQOTSQNIK-NQLAEALNATNIYVLDKIKLNAMNIIIRDKRPYD- 599
DB 4258 YLTATLDVLPN-DKVNIVNISFNKPSIAEN-----IGDKSNII-----YNNDS 4304

QY 600 -----RNIIAVGA---DESVVKEAHREVINGSTEGLLNIDKIRKILSGVIEIETDE 650
DB 4305 IPKLEINNDIIVNGPINKIEIVVKNQOK--NNIDVDLGLQINPKIAHNL-FAIAFKSTN 4361

QY 651 GLKEVIND--RYDMLNISSL-RQDGKTFIDF-----KKY----- 681
DB 4362 -----NDIETVINGSSLVNDGKTSIRPTLNNLKANKLYSLVDVYLVNNSNTIVE 4415
Hypothetical protein UU482.

Db 1922 EDIIQERFHNHNIKYENKNCYSDCDNIYKINYDELNDSTQTKELNEGKSNNGKAEAW 1981
Qy 106 - IREKGLGRLY-QIKIOYQRENPTKGLDFKLYWTDSONKEVEISSDNLQLPKQKSSNS 163
Db 1982 IIDIKNNETYPIKIDKEKNEDEK--KNKYMKYKNDKNIKMGSSN--KSMKKNKNS 2036
Qy 164 RKKSTSGAPVDPORDNDGIPDSLEVEGYTVDDVKNKFTFLSPWISNHEKGLTKYKSSP 223
Db 2037 NKMKH-----IFLSVNNKGY-----NKSINKKYENNIN-----KONK 2069
Qy 224 EKWSTASDPYGF--EKVTGRI-DQNVSPPEARHPLVAAYPIVHVHME-----NILS 272
Db 2070 DKLNLVNSIKLVQSKIKQELSNKNSKD-----ILNFEITKIKKSKKETKNTNTNN 2124
Qy 273 KNEOSTONTSETPTISKNVSTSRHTSEV-HGNAEVHASFPLIGSVSG-PSNSNS 330
Db 2125 KUNDNNNDNNINFERINDNINSYVNLNKKVPYDNINI-----DEGNYINYGPYAPDGS 2179
Qy 331 TVAIHSLSLAGERWAEWTGLNTA-----DTARLNA-----NIRVNTGTAPIYVNL 378
Db 2180 TI-----YVWVNIDTNYMYKVFDSKONINHHVLLNVPYLN-----NDL 2221
Qy 379 PTTSLVLGKNQFLATIKAKENOLSO-----ILAPNYPYKSLAPIALNAQKDFSTPTIM 434
Db 2222 LNNVILNQNMMNLENINTIGSVQPFVTCDFY--ANNIKSYLD-----PNLQ 2271
Qy 435 NYNQLELEKTKQLRLDT-----DOVYGNIAFYNFENGVRVDT-----GSN 476
Db 2272 NNNYFNDMQLNHNLNNNLSINVIDQNYLSYNNLNCINGNINKDTCKDIVIGPNNIN 2331
Qy 477 WSEVLPQIQTETARIIFNGKDLNVERRIAANVPDPLE-----TTKPDMLKEALK 528
Db 2332 QNGI-PTI-ELDDTILKN--DVNLILNNNVNHSVYVEMLNNIQNVAKLYNDIQENLH 2387
Qy 529 IAFGP--NEPNGLOYQKGD-----ITFEDFNDOQTSONIKNQLAELNATNIYTV 577
Db 2388 ITNSLNNNNNNNNMLGGDLVNSYLFNINSFNPNWYIYNNNNNNNNNNCINYL 2447
Qy 578 LDKIKLNAK-----MNLILDKRFHYDRNNAIVGADESVVKAHREVINSSTEGLLNID 632
Db 2448 LDRKEINVKEEIMNMSVLEDTNKEGND-----MKDNKYIINK--ENNVSQND 2496
Qy 633 KDRKLSGYVIEIEETGLKEVINDRYDMLISSLRQDKFTFIDPKYNDKLPYISNP 692
Db 2497 NNIDIIISNGQNIHNTANTWKEKRNRLRDDDIYKQMKRNSLDFPKLDSE-----KK 2550
Qy 693 NYKVV--YAVTKENT--IINPSENGDTSTNGIKKILIFSCKGYE 733
Db 2551 NKSNIIEGARKKMTNTQLLNYSEN--RKGSLDMSTYADKYLE 2593
RESULT 15
Q8RJN9 PRELIMINARY; PRT; 1404 AA.
AC Q8RJN9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Variable membrane protein precursor.
GN VMP.
OS Mycoplasma hominis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2098;
RN [1]
RC STRAIN=132;
RA Boesen T.;
RT "Gene and Protein Structure of the Mycoplasma hominis Vaa adhesin.";
RL Thesis (2001), Department of Department of Medical Microbiology and.
DR EMBL; AJ416752; CAC95143.1; --
DR InterPro; IPR006864; LMP.
DR Pfam; PF04776; LMP; 3.

KW Signal. 1 24 POTENTIAL.
FT SIGNAL 25 1404 VARIABLE MEMBRANE PROTEIN.
SQ SEQUENCE 1404 AA; 160137 MW; 2C03F666B3473CAB CRC64;
Query Match 5.1%; Score 194; DB 2; Length 1404;
Best Local Similarity 19.7%; Pred. NO. 0.5;
Matches 158; Conservative 124; Mismatches 282; Indels 240; Gaps 34;
Qy 47 SSELNIPSENYFQSAISWISGFIKYSDEYTFATSDHNVHTMWVDDQ-----EVIN 98
Db 312 NSSIEDIQNTKDIKAIES-----LTNKINDQKQKMLNEVIN 351
Qy 99 KASNSNKIRLEK-GRLYQIKIQORE-----NPTKEGLDFKLYWTDSON 141
Db 352 KAKELVKLVDSDEIQAKTQDQEIQKASQVWASNDTKAINSSKTSIDAKI--TDITK 409
Qy 142 KKEVI--SSNLIQLPKQKSSNSRK--KSTSGAPVDPORDNDGIPDSLEVEGYTVDVKN 198
Db 410 KLEAFNATKXLEFTKLOETRSNIDKPLTPEVKANPYITLVNE-----LEVAK 457
Qy 199 KRTFLSPWISNHEKGLTKYKSS-PEKWTASDPSYDFEKTG-----RI 243
Db 458 K---AKEWSESSNSKSDIVAANNEKQAFQTAQSSKNDAKTSNEBKAKLSASNAKEL 514
Qy 244 DKVVS-----PEARHPLVAAYPIVHVDMENILSKNQDSTQNTDSETRTISKNTSTSR 297
Db 515 DKNLTDSDGEIQOAKAELAE-----EVEKANQAITSNNTKEIQN--SNTSLNK-----561
Qy 298 THTSEVHG-----NAEVHASFFDIGSVSA-----GFSNSNSTVAIDHSLSLAGERTWA 347
Db 562 --LSEVKNKLDKFSNEKEAEFNKLEASRAIKFEFINENNTNPNYTAIQLKQAKLDKANS 619
Qy 348 ETWGLNTADTARLNAIR-----YVNTGTAPIYVNLPTT--SLVAGKN-----388
Db 620 ITKSSNKSDDIATNQALQEAIAIEKTEKESANSQAVKNTLNLETIGKAKELDKNLTDS 679
Qy 389 ----OTLATIKAKENOLSOILAPNYPYKSLAPIALNAQKDFSTPTITMNYNQFLELEK 444
Db 680 GETQOAKAELTNEIKANQTIASNTALMENSNTSLNKISEVQNKLDRFNNDKKAEPNK 739
Qy 445 TKQLRLDT-----QVYGNIAFYN-FENGVRVDTGSNWS 478
Db 740 LQELKNKIDDFEKKKNKNEIYSKFNLDLDELINKSVQIKNSIGSNSSNKKDIIVANKOMQ 799
Qy 479 EVLPQIQTETARI-----IFN--GKDLNVERRIAANVPSPLETTPKPDMLKEALKI 529
Db 800 DALNELQAKWAEIHUKTFQEFNEHKELENLIKKEDAKEVGTDEANTAITNNDVKENSSI 859
Qy 530 AFGFNEPNGNLOVQGDITBDFNFDQOTS---QNIKNQLA-ELNATNIYTVLDKIKLNA 585
Db 860 -----BEITKATKALDEAKSKLDQKINTQKATELSLN-----ESKE 896
Qy 586 KQNTILI-----RDKRPHYDRNNAIVGADES--VKEAHEVINSSTEGLLN 630
Db 897 KLNLIITSSSNQVSAEISKAKVLEBINLNSLNDSSIKSLKEATQIKDAETQ-LTKE 955
Qy 631 IDK-----DIRKILSGYIYE-----TEDTEGLKEVINDRYDML 663
Db 956 IEKAKIEKTDKLRKFNEVKKSLEDLIKDDAIOVGTDDAPKLEDDNNNINE--NSSIEEII 1014
Qy 664 NISSLRQDKFTFIDPKYNDKLP 687
Db 1015 NATKLEDGSKLDKIKTKKQPL 1038

Search completed: May 3, 2004, 19:39:56
Job time : 39.6261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:26:27 ; Search time 49.1208 Seconds
(without alignments)
4227.791 Million cell updates/sec

US-09-848-909A-10

Title: Perfect score: 3774

Sequence: 1 EVKQENRLNESBSSQGLL.....TSTNGIKKILFSSKKGYBIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1980s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3774	100.0	735	5 AAMS1493	Aam51493 Anthrax P
2	3768	99.8	735	5 AAMS1485	Aam51485 Anthrax P
3	3767	99.8	735	5 AAMS1491	Aam51491 Anthrax P
4	3766	99.8	735	5 AAMS1495	Aam51495 Anthrax P
5	3763	99.7	735	5 AAMS1490	Aam51490 Anthrax P
6	3762	99.7	735	5 AAMS1487	Aam51487 Anthrax P
7	3762	99.7	735	5 AAMS1489	Aam51489 Anthrax P
8	3761	99.7	735	5 AAR60179	Aar60179 Protective
9	3761	99.7	735	5 AAMS1483	Aam51483 Anthrax P
10	3761	99.7	735	5 AAMS1488	Aam51488 Anthrax P
11	3761	99.7	735	5 AAMS1494	Aam51494 Anthrax P
12	3761	99.7	736	3 AAY56959	Aay56959 B. anthra
13	3761	99.7	763	3 AAY56960	Aay56960 B. anthra
14	3761	99.7	764	3 AAY56958	Aay56958 B. anthra
15	3761	99.7	764	4 AAB47306	Aab47306 Wild type
16	3761	99.7	857	7 ADE65872	Ade65872 Bacillus
17	3760	99.6	735	5 AAMS1484	Aam51484 Anthrax P
18	3759	99.6	735	5 AAMS1486	Aam51486 Anthrax P
19	3758	99.6	735	5 AAE18289	Aae18289 Bacillus
20	3758	99.6	764	6 AAE35717	Aae35717 Bacillus
21	3753	99.4	735	5 AAMS1492	Aam51492 Anthrax P
22	3752	99.4	735	5 AAMS1500	Aam51500 Anthrax P
23	3752	99.4	735	5 AAMS1499	Aam51499 Anthrax P
24	3746	99.3	764	6 ASP71693	Abp71693 B. anthra
25	3716.5	98.5	903	2 AAR60183	Aar60183 PA(1-725)

RESULT 1

AAMS1493
ID AAMS1493 standard; protein; 735 AA.
XX
AC AAMS1493;
XX
DT 01-FEB-2002 (first entry)
XX
DE Anthrax PA mutant X397D/D425K.
XX
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mtein.
XX
OS Bacillus anthracis.
XX
FH Key Location/Qualifiers
FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"
FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"
XX
PN WO200182788-A2.
XX
PD 08-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014372.
XX
PR 04-MAY-2000; 2000US-0201800P.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Collier RJ, Sellman BR;
XX
DR WPI; 2002-017725/02.
XX
PT Protecting humans against anthrax using mutant B groups (anthrax
protective antigens) of the pore-forming binary A-B toxin of Bacillus
anthracis.
XX
PS Claim 4; Page; 77pp; English.
XX
CC The invention relates to antibacterial agents comprising mutant forms of
pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants
in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
moiety is anthrax protective antigen (PA) and using these mutants or
compositions of them for protecting against Bacillus anthracis infections
in humans, especially as vaccines. Note: The present sequence is not

ALIGNMENTS

26	3620	95.9	711	5	AAMS1498	Aam51498 Anthrax P
27	3615	95.8	711	5	AAMS2113	Aam52113 Anthrax P
28	3614	95.8	711	5	AAMS1497	Aam51497 Anthrax P
29	3607	95.6	711	5	AAMS1496	Aam51496 Anthrax P
30	3584	95.0	719	2	AAR60193	Aar60193 Modified
31	3045	80.7	595	3	Aae18288	Aae18288 Bacillus
32	2906	77.0	569	3	AAY56961	Aay56961 B. anthra
33	2496	66.1	487	5	Aae18285	Aae18285 Bacillus
34	2181	57.8	426	5	Aae18287	Aae18287 Bacillus
35	1632	43.2	318	5	Aae18286	Aae18286 Bacillus
36	1338	35.5	258	5	Aae18284	Aae18284 Bacillus
37	1275	33.8	288	5	AAMS0707	Aam50707 Bacillus
38	847.5	22.5	1052	4	Aae07903	Aae07903 C. botuli
39	846	22.4	1032	4	Aae07901	Aae07901 C. botuli
40	842	22.3	1092	4	Aae07900	Aae07900 C. botuli
41	833.5	22.1	1112	4	Aae07902	Aae07902 C. botuli
42	818	21.7	721	6	Aae35719	Aae35719 Clostridi
43	775.5	20.5	880	2	AAM60224	Aaw60224 Bacillus
44	775	20.5	881	3	AAY59277	Aay59277 MIS toxin
45	774	20.5	884	2	AAR91239	Aar91239 B. cereus

CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
SQ Sequence 735 AA;

Query Match 100.0%; Score 3774; DB 5; Length 735;
Best Local Similarity 100.0%; Pred. No. 4.3e-241;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLLYGFFSDFLNFOAPMVVTSSTGDLSPSELENIPISENQYF 60
Db 1 EVKQENRLNESSESSQGLLYGFFSDFLNFOAPMVVTSSTGDLSPSELENIPISENQYF 60

QY 61 QSAIWSGFIKVKSDSEYTFATADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240

QY 241 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 300
Db 241 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 300

QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYNNVLPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVLPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITWYNNQFLELTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITWYNNQFLELTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480

Query Match 100.0%; Score 3774; DB 5; Length 735;
Best Local Similarity 100.0%; Pred. No. 4.3e-241;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLLYGFFSDFLNFOAPMVVTSSTGDLSPSELENIPISENQYF 60
Db 1 EVKQENRLNESSESSQGLLYGFFSDFLNFOAPMVVTSSTGDLSPSELENIPISENQYF 60

QY 61 QSAIWSGFIKVKSDSEYTFATADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDSEYTFATADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240

QY 241 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 300
Db 241 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 300

QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYNNVLPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVLPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITWYNNQFLELTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITWYNNQFLELTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480

QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLTKALIAFGFNPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMLTKALIAFGFNPNGNL 540

QY 541 QYQCKDITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKKNLILIRDKRPHYDR 600
Db 541 QYQCKDITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKKNLILIRDKRPHYDR 600

QY 601 NNIAVGADESVMKEAHEVINSSTEGILLNIDKIDKISGYIVEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGILLNIDKIDKISGYIVEIDTEGLKEVINDRY 660

QY 661 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 2
AAM51485
ID AAM51485 standard; protein; 735 AA.
XX
AC AAM51485;
XX
DT 01-FEB-2002 (first entry)
XX
DE Anthrax PA mutant K397D.
XX
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

Db 421 LNAQDFSSPTITMNYNQFLEKTKQLRLDQVIGNIATYFENGVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Qy 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNLIIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNLIIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
Qy 661 DMLNISRQDGKTFIDFKYNDKPLIYISPNYKVNVAVTKENTLIINPSENGDTSTNG 720
Db 661 DMLNISRQDGKTFIDFKYNDKPLIYISPNYKVNVAVTKENTLIINPSENGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 3
AAM51491
ID AAM51491 standard; protein; 735 AA.
AC AAM51491;
XX
DT 01-FEB-2002 (first entry)
DE Anthrax PA mutant D425K.
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX
OS Bacillus anthracis.
XX Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 425
FT /note= "Wild-type Asp substituted by Lys"
XX
XX WO200182788-A2.
XX
XX 08-NOV-2001.
XX
XX 04-MAY-2001; 2001WC-US014372.
XX
XX 04-MAY-2000; 2000US-0201800P.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX Collier RJ, Sellman BR;
XX
XX WPI; 2002-017725/02.
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
PT anthracis.
XX
XX Claim 4; Page; 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAM51483)
XX

SQ Sequence 735 AA;
Query Match 99.8%; Score 3767; DB 5; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.2e-240;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLLNESSESSQGLLYFSDLNFAQPMVVTSTTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLLNESSESSQGLLYFSDLNFAQPMVVTSTTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWGSGFIKVKSDSEYTFATSDAHNVMTWDDQVINKASNSKIRLEKRLYQIKY 120
Db 61 QSAIWGSGFIKVKSDSEYTFATSDAHNVMTWDDQVINKASNSKIRLEKRLYQIKY 120
Qy 121 QENPTEKGLDFKLYWTDSONKKEVISSDNQLQPELKOKSNSRKRSTAGPTVPDRDN 180
Db 121 QENPTEKGLDFKLYWTDSONKKEVISSDNQLQPELKOKSNSRKRSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVQKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVQKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240
Qy 241 GRIDKNVSPERHPLVAAPIVHVDMENIILSKNDQSTQNTDSTRTISKTSTSRHT 300
Db 241 GRIDKNVSPERHPLVAAPIVHVDMENIILSKNDQSTQNTDSTRTISKTSTSRHT 300
Qy 301 SEVHGNAEVHAGFFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAGFFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSTPTITMNYNQFLEKTKQLRLDQVYIGNIATYFENGVRVDTGSNWSEV 480
Db 421 LNAQKDFSTPTITMNYNQFLEKTKQLRLDQVYIGNIATYFENGVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Qy 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNLIIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNLIIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
Qy 661 DMLNISRQDGKTFIDFKYNDKPLIYISPNYKVNVAVTKENTLIINPSENGDTSTNG 720
Db 661 DMLNISRQDGKTFIDFKYNDKPLIYISPNYKVNVAVTKENTLIINPSENGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 4
AAM51495
ID AAM51495 standard; protein; 735 AA.
XX
XX AC AAM51495;
XX
XX DT 01-FEB-2002 (first entry)
XX Anthrax PA mutant K397D/D425K/F427A.
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX
XX OS Bacillus anthracis.

OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"
 FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 427 /note= "Wild-type Phe substituted by Ala"
 FT
 XX
 PN WO200182788-A2.
 XX
 XX
 XX 08-NOV-2001.
 XX
 XX 04-MAY-2001; 2001WO-US014372.
 XX
 XX 04-MAY-2000; 2000US-0201800P.
 XX
 XX (HARD) HARVARD COLLEGE.
 XX
 XX Collier RJ, Sellman BR;
 XX
 XX WPI; 2002-017725/02.
 XX
 XX Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.
 XX
 XX Claim 4; Page; 77pp; English.
 XX
 XX The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections
 CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)
 XX
 XX Sequence 735 AA;
 SQ

Query Match 99.8%; Score 3766; DB 5; Length 735;
 Best Local Similarity 99.9%; Pred. No. 1.5e-240;
 Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSSTTGDLSTPSSSELENIPSENOYF 60
 DB 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSSTTGDLSTPSSSELENIPSENOYF 60
 QY 61 QSAIWSGFIKVKKSDEYTFATGADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
 DB 61 QSAIWSGFIKVKKSDEYTFATGADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
 QY 121 QRENPTKEGLDFKLYWTDSONKEVVISDNLOLPELKQKSSNRKKGSTASGPTVPDRDN 180
 DB 121 QRENPTKEGLDFKLYWTDSONKEVVISDNLOLPELKQKSSNRKKGSTASGPTVPDRDN 180
 QY 181 DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKMWSTASDPYDFEYVT 240
 DB 181 DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKMWSTASDPYDFEYVT 240
 QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
 DB 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
 QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
 DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
 QY 361 NANIRYVNTGTAPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYV 420
 DB 361 NANIRYVNTGTAPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYVLPYV 420

QY 421 LNAQKDFSTPITMNYNQFLELEKTKQLRLDQVYGNIAFYNFENGVRVYDTGNNWSEV 480
 DB 421 LNAQKDFSTPITMNYNQFLELEKTKQLRLDQVYGNIAFYNFENGVRVYDTGNNWSEV 480
 QY 481 LPQIQTETARIIFNGKDLNLVERRITAAVNSPDPLETTKPDMTLKEALKIAFGNEPENG 540
 DB 481 LPQIQTETARIIFNGKDLNLVERRITAAVNSPDPLETTKPDMTLKEALKIAFGNEPENG 540
 QY 541 QYQGGKDIETFDNFQDQTSQNIKNQLAELNATNIYTVLDKILNAKMNILIRDKRFHYDR 600
 DB 541 QYQGGKDIETFDNFQDQTSQNIKNQLAELNATNIYTVLDKILNAKMNILIRDKRFHYDR 600
 QY 601 NNIAVGADESUVKEAHREVINSSTGLLLNDKDKIRKILSGYVIEIDTEGKEVINDRY 660
 DB 601 NNIAVGADESUVKEAHREVINSSTGLLLNDKDKIRKILSGYVIEIDTEGKEVINDRY 660
 QY 661 DMLNLSLQDQGTIDFKYNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720
 DB 661 DMLNLSLQDQGTIDFKYNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720
 QY 721 IKKILIFSKGYEIG 735
 DB 721 IKKILIFSKGYEIG 735
 RESULT 5
 AAM51490
 ID AAM51490 standard; protein; 735 AA.
 XX
 XX AAM51490;
 XX
 XX 01-FEB-2002 (first entry)
 DT Anthrax PA mutant D425E.
 DE
 XX
 XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.
 KW
 XX Bacillus anthracis.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 425 /note= "Wild-type Asp substituted by Glu"
 FT
 FT
 XX WO200182788-A2.
 PN
 XX
 XX 08-NOV-2001.
 PD
 XX
 XX 04-MAY-2001; 2001WO-US014372.
 XX
 XX 04-MAY-2000; 2000US-0201800P.
 XX
 XX (HARD) HARVARD COLLEGE.
 XX
 XX Collier RJ, Sellman BR;
 XX
 XX WPI; 2002-017725/02.
 XX
 XX Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.
 XX
 XX Claim 4; Page; 77pp; English.
 XX
 XX The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections
 CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)
 CC

XX	SQ	Sequence 735 AA;	
XX		Query Match 99.7%; Score 3763; DB 5; Length 735;	
XX		Best Local Similarity 99.7%; Pred. No. 2.3e-240;	
XX		Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
QY	1	EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSPSSELENIPSENQYF 60	
DB	1	EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSPSSELENIPSENQYF 60	
QY	61	QSAIWSGFIKVKSDGYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120	
DB	61	QSAIWSGFIKVKSDGYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120	
QY	121	QRENPTKGLDFKLYWTDSONKKEVSSDNILQLPELKQSSNSRKKRSTAGPTVPDRN 180	
DB	121	QRENPTKGLDFKLYWTDSONKKEVSSDNILQLPELKQSSNSRKKRSTAGPTVPDRN 180	
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLKYKSSPEKWTASDPYDFEKT 240	
DB	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLKYKSSPEKWTASDPYDFEKT 240	
QY	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300	
DB	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300	
QY	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
DB	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
QY	361	NANIRYVNTGTAPIYNNVLTSLVGLKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420	
DB	361	NANIRYVNTGTAPIYNNVLTSLVGLKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420	
QY	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNNWSEV 480	
DB	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNNWSEV 480	
QY	481	LPOIQTETARIIPNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPQNL 540	
DB	481	LPOIQTETARIIPNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPQNL 540	
QY	541	QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRPHYDR 600	
DB	541	QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRPHYDR 600	
QY	601	NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660	
DB	601	NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660	
QY	661	DMNLSLRQDGKTFIDFKKYNKPLPYSNPNYKNVYAVTKENTIINPSENGDTSTNG 720	
DB	661	DMNLSLRQDGKTFIDFKKYNKPLPYSNPNYKNVYAVTKENTIINPSENGDTSTNG 720	
QY	721	IKKILFSKKGVEIG 735	
DB	721	IKKILFSKKGVEIG 735	
XX		RESULT 6	
XX		AAMS1487	
XX		ID AAMS1487 standard; protein; 735 AA.	
XX		AC AAMS1487;	
XX		DT 01-FEB-2002 (first entry)	
XX		DE Anthrax PA mutant K397Q.	
XX		XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;	
XX		KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; muten.	
XX		XX	

OS	Bacillus anthracis.		
OS	Synthetic.		
XX			
XX	Key	Location/Qualifiers	
XX	Misc-difference 397	/note= "Wild-type Lys substituted by Gln"	
XX			
XX	WO200182788-A2.		
XX	08-NOV-2001.		
XX	04-MAY-2001; 2001WO-US014372.		
XX	04-MAY-2000; 2000US-0201800P.		
XX	(HARD) HARVARD COLLEGE.		
XX	Collier RJ, Sellman BR;		
XX	WPI; 2002-017725/02.		
XX	Protecting humans against anthrax using mutant B groups (anthrax		
XX	protective antigens) of the pore-forming binary A-B toxin of Bacillus		
XX	anthracis.		
XX	Claim 4; Page; 77pp; English.		
XX	The invention relates to antibacterial agents comprising mutant forms of		
XX	pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants		
XX	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B		
XX	moiety is anthrax protective antigen (PA) and using these mutants or		
XX	compositions of them for protecting against Bacillus anthracis infections		
XX	in humans, especially as vaccines. Note: The present sequence is not		
XX	given in the specification but is derived from the Bacillus anthracis		
XX	wild-type PA protein sequence shown in figure 13 (AAMS1483)		
XX	Sequence 735 AA;		
XX	Query Match 99.7%; Score 3762; DB 5; Length 735;		
XX	Best Local Similarity 99.7%; Pred. No. 2.7e-240;		
XX	Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY	1	EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSPSSELENIPSENQYF 60	
DB	1	EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSPSSELENIPSENQYF 60	
QY	61	QSAIWSGFIKVKSDGYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120	
DB	61	QSAIWSGFIKVKSDGYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120	
QY	121	QRENPTKGLDFKLYWTDSONKKEVSSDNILQLPELKQSSNSRKKRSTAGPTVPDRN 180	
DB	121	QRENPTKGLDFKLYWTDSONKKEVSSDNILQLPELKQSSNSRKKRSTAGPTVPDRN 180	
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLKYKSSPEKWTASDPYDFEKT 240	
DB	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLKYKSSPEKWTASDPYDFEKT 240	
QY	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300	
DB	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300	
QY	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
DB	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
QY	361	NANIRYVNTGTAPIYNNVLTSLVGLKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420	
DB	361	NANIRYVNTGTAPIYNNVLTSLVGLKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420	
QY	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNNWSEV 480	
DB	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNNWSEV 480	

481 LPQIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540
 481 LPQIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540
 541 QYQKDIETEDFNPDQTSQNIKQVLAELNATNIYTVLDKIKNKAKMILIRDKGFHYDR 600
 541 QYQKDIETEDFNPDQTSQNIKQVLAELNATNIYTVLDKIKNKAKMILIRDKGFHYDR 600
 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 661 DMLNLSLRQDGKTFDFKXNDKLPYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720
 661 DMLNLSLRQDGKTFDFKXNDKLPYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720
 721 IKKILIFSCKGYEIG 735
 721 IKKILIFSCKGYEIG 735
 RESULT 7
 AAM51489
 ID AAM51489 standard; protein; 735 AA.
 AC
 AC
 DT 01-FEB-2002 (first entry)
 DE Anthrax PA mutant D425N.
 XX
 XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutuin.
 XX
 XX Bacillus anthracis.
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 425
 FT /note= "Wild-type Asp substituted by Asn"
 XX
 XX
 FN WO200132788-A2.
 XX
 XX 08-NOV-2001.
 XX
 XX 04-MAY-2001; 2001WO-US014372.
 XX
 XX 04-MAY-2000; 2000US-0201800P.
 XX
 XX (HARD) HARVARD COLLEGE.
 XX
 XX Collier RJ, Sellman BR;
 XX
 XX WPI; 2002-017725/02.
 XX
 XX Protecting humans against anthrax using mutant B groups (anthrax
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
 PT anthracis.
 XX
 XX Claim 4; Page; 77pp; English.
 XX
 CC The invention relates to antibacterial agents comprising mutant forms of
 CC pore-forming toxins (AAM5113 and AAM51484-AAM51500), especially mutants
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
 CC moiety is anthrax protective antigen (PA) and using these mutants or
 CC compositions of them for protecting against Bacillus anthracis infections
 CC in humans, especially as vaccines. Note: The present sequence is not
 CC given in the specification but is derived from the Bacillus anthracis
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)
 XX
 XX Sequence 735 AA;
 SQ

Query Match 99.7%; Score 3762; DB 5; Length 735;
 Best Local Similarity 99.7%; Pred. No. 2.7e-240;
 Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EVKQENRLLNESSESSQGLLGYYPSDLNFQAPMVVTSSTTCGLSIPSELEINISENQYF 60
 DB 1 EVKQENRLLNESSESSQGLLGYYPSDLNFQAPMVVTSSTTCGLSIPSELEINISENQYF 60
 QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMVWDDQEVINKASNNKIRLEKGRLYQIKIY 120
 DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMVWDDQEVINKASNNKIRLEKGRLYQIKIY 120
 QY 121 QRENTEKGLDKPLKWTDSQNKKEVSSDNLQLOLPELKOKSSNKKSTAGTPVPDRN 180
 DB 121 QRENTEKGLDKPLKWTDSQNKKEVSSDNLQLOLPELKOKSSNKKSTAGTPVPDRN 180
 QY 181 DGIPOSLEVEGYTVDVQNKRTPLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
 DB 181 DGIPOSLEVEGYTVDVQNKRTPLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
 QY 241 GRIDKNVSPERARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRISKTSTHT 300
 DB 241 GRIDKNVSPERARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRISKTSTHT 300
 QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLTADTARL 360
 DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLTADTARL 360
 QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATKADENQLSOILAPNNYPSKNLAPIA 420
 DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATKADENQLSOILAPNNYPSKNLAPIA 420
 QY 421 LNAQKDFSTPTIMYNOFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGNNSEV 480
 DB 421 LNAQKDFSTPTIMYNOFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGNNSEV 480
 QY 481 LPOIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540
 DB 481 LPOIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540
 QY 541 QYQKDIETEDFNPDQTSQNIKQVLAELNATNIYTVLDKIKNKAKMILIRDKGFHYDR 600
 DB 541 QYQKDIETEDFNPDQTSQNIKQVLAELNATNIYTVLDKIKNKAKMILIRDKGFHYDR 600
 QY 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 DB 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
 QY 661 DMLNLSLRQDGKTFDFKXNDKLPYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720
 DB 661 DMLNLSLRQDGKTFDFKXNDKLPYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720
 QY 721 IKKILIFSCKGYEIG 735
 DB 721 IKKILIFSCKGYEIG 735
 RESULT 8
 AAR60179
 ID AAR60179 standard; protein; 735 AA.
 XX
 XX
 AC AAR60179;
 XX
 XX 25-MAR-2003 (revised)
 DT 03-APR-1995 (first entry)
 XX
 XX Protective antigen of Bacillus anthracis.
 XX Anthrax; Bacillus anthracis; fusion protein; protective antigen;
 KW protective antigen; cell killing; targeting; targeting; pathogen;
 KW intracellular; HIV; human immunodeficiency virus; toxin.
 XX
 XX Bacillus anthracis.
 OS

XX WO9418332-A2.
FN XX
PD XX
XX 18-AUG-1994.
XX 14-FEB-1994; 94WO-US001624.
XX 12-FEB-1993; 93US-00021601.
PR 25-JUN-1993; 93US-00082849.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Leppla SH, Klimpel K, Arora N, Singh Y, Nichols PJ;
XX WPI; 1994-279753/34.
DR N-PSDB; AAQ70180.
XX Nucleic acid encoding anthrax toxin fusion protein - useful for
PT targeting toxin to specific cells, eg for killing tumour cells or HIV-
PT infected cells.
XX Disclosure; Page 81-83; 124pp; English.
XX The sequence encoding the protective antigen of Bacillus anthracis may be
CC used in the construction of a nucleic acid which encodes a fusion protein
CC comprising the anthrax protective antigen binding domain of the native
CC anthrax lethal factor and a sequence encoding an activity inducing domain
CC of a second protein. The fusion proteins are useful for the specific
CC killing of tumour cells or the killing of cells infected with
CC intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to
CC correct FN field.)
XX SQ Sequence 735 AA;
Query Match 99.7%; Score 3761; DB 2; Length 735;
Best Local Similarity 99.7%; Pred. No. 3.1e-240;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60
QY 61 QSAIWSGFIKVKKSDYEYFATADNHVTWVDDQEVINKAGNSNKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYEYFATADNHVTWVDDQEVINKAGNSNKIRLEKRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSNRSKRSTAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSNRSKRSTAGTPVDRDN 180
QY 181 DGIPDSLEVEGYTVQVKNKRTFLSPWISNHEKGLTKYKSSPEKXSTASDPYSDFEVT 240
Db 181 DGIPDSLEVEGYTVQVKNKRTFLSPWISNHEKGLTKYKSSPEKXSTASDPYSDFEVT 240
QY 241 GRIDKNVSPPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNVTSRTH 300
Db 241 GRIDKNVSPPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNVTSRTH 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NAMRYVNTGTAPIYVLPVTSVLGKNOTLATIKADENQLSQILAPNNYFSPKNLAPIA 420
Db 361 NAMRYVNTGTAPIYVLPVTSVLGKNOTLATIKADENQLSQILAPNNYFSPKNLAPIA 420
QY 421 LNAQKQFSSPTITWYNNQFLESEKTKQLRLDTPQVYGNATYFNFGNRYVDTGNSWSEV 480
Db 421 LNAQKQFSSPTITWYNNQFLESEKTKQLRLDTPQVYGNATYFNFGNRYVDTGNSWSEV 480
QY 481 LPQIQETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNENGNL 540
Db 481 LPQIQETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNENGNL 540

QY 541 QYQKDIETEDFNFEDQOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRFHYDR 600
Db 541 QYQKDIETEDFNFEDQOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRFHYDR 600
QY 601 NNIAGADESVVKEAHREVINSSTEGILLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGILLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDPKYNDKLPYISNPYKAVYAVTKENTINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDPKYNDKLPYISNPYKAVYAVTKENTINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
RESULT 9
AAM51483
ID AAM51483 standard; protein; 735 AA.
XX AC AAM51483;
XX 01-FEB-2002 (first entry)
DT Anthrax PA protein.
DE Anthrax PA protein.
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.
XX Bacillus anthracis.
XX WO200182788-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US014372.
XX 04-MAY-2000; 2000US-0201800P.
XX (HARD) HARVARD COLLEGE.
XX Collier RJ, Sellman BR;
PI WPI; 2002-017725/02.
XX N-PSDB; AAI99904.
DR Protecting humans against anthrax using mutant B groups (anthrax
PT protective antigens/ of the pore-forming binary A-B toxin of Bacillus
PT anthracis.
XX Disclosure; Fig 13; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. The present sequence is that of the
CC anthrax PA protein
XX SQ Sequence 735 AA;
Query Match 99.7%; Score 3761; DB 5; Length 735;
Best Local Similarity 99.7%; Pred. No. 3.1e-240;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60
QY 61 QSAIWSGFIKVKKSDYEYFATADNHVTWVDDQEVINKAGNSNKIRLEKRLYQIKIY 120

```

Db 61 QSAIWSGFIRKVKSDYFTFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDPLKLYWTDSONKKEVYSSDNLQIPKOKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDPLKLYWTDSONKKEVYSSDNLQIPKOKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPLDSLEVEGYTVDVNKNKRTFLSPWISNIHEKKGLTKYKSPKMWSTASDPYDFEYV 240
Db 181 DGIPLDSLEVEGYTVDVNKNKRTFLSPWISNIHEKKGLTKYKSPKMWSTASDPYDFEYV 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSEPTISKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSEPTISKNSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVDTSNWSSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVDTSNWSSEV 480
QY 481 LPOIQTETARIIFNGKOLNVERRIAAVNPSDPLETTKPDMTLKEALKIARFNPNGNL 540
Db 481 LPOIQTETARIIFNGKOLNVERRIAAVNPSDPLETTKPDMTLKEALKIARFNPNGNL 540
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAHREVINSTEGLLNIDKIDKILSGYIIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHREVINSTEGLLNIDKIDKILSGYIIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRQDKTFIDPKYNDKPLIYISNPNKYVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNISSLRQDKTFIDPKYNDKPLIYISNPNKYVAVTKENTIIINPSNGDTSTNG 720
QY 721 IKKILIFSKGGEIG 735
Db 721 IKKILIFSKGGEIG 735

RESULT 10
AAW51488
ID AAW51488 standard; protein; 735 AA.
XX AC AAW51488;
XX DT 01-FEB-2002 (first entry)
XX DE Anthrax PA mutant D425A.
XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX OS Bacillus anthracis.
XX OS Synthetic.
XX FT Key Location/Qualifiers
XX FT Misc-difference 425
XX FT /note= "Wild-type Asp substituted by Ala"
XX PN WO200182788-A2.
XX PD 08-NOV-2001.
XX PF 04-MAY-2001; 2001WO-US014372.

```

```

XX 04-MAY-2000; 2000US-0201800P.
XX (HARD ) HARVARD COLLEGE.
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX Claim 4; Page; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAW52113 and AAW51484-AAW51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAW51483)
XX SQ Sequence 735 AA;

Query Match          99.7%; Score 3761; DB 5; Length 735;
Best Local Similarity 99.7%; Pred No 3 1e-240;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYYFSDLNFPQPMVVTSTTGDLSPSELENISENQYF 60
Db 1 EVKQENRLNSESQGLLYYFSDLNFPQPMVVTSTTGDLSPSELENISENQYF 60
QY 61 QSAIWSGFIRKVKSDYFTFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIRKVKSDYFTFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDPLKLYWTDSONKKEVYSSDNLQIPKOKSSNSRKRSTASGTPVDRDN 180
Db 121 QRENPTKGLDPLKLYWTDSONKKEVYSSDNLQIPKOKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPLDSLEVEGYTVDVNKNKRTFLSPWISNIHEKKGLTKYKSPKMWSTASDPYDFEYV 240
Db 181 DGIPLDSLEVEGYTVDVNKNKRTFLSPWISNIHEKKGLTKYKSPKMWSTASDPYDFEYV 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSEPTISKNSTSRHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSEPTISKNSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVDTSNWSSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVDTSNWSSEV 480
QY 481 LPOIQTETARIIFNGKOLNVERRIAAVNPSDPLETTKPDMTLKEALKIARFNPNGNL 540
Db 481 LPOIQTETARIIFNGKOLNVERRIAAVNPSDPLETTKPDMTLKEALKIARFNPNGNL 540
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLRDKRFHYDR 600
QY 601 NNIAVGADESVMKEAHREVINSTEGLLNIDKIDKILSGYIIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHREVINSTEGLLNIDKIDKILSGYIIEDETEGLKEVINDRY 660

```

QY	661	DMLNISSLRODQKTFIDFKYNDKLPYISNPYKVNVAVTYKNTIINPSENGDTSTNG	720
Db	661	DMLNISSLRODQKTFIDFKYNDKLPYISNPYKVNVAVTYKNTIINPSENGDTSTNG	720
QY	721	IKKILIFSKKGYEIG	735
Db	721	IKKILIFSKKGYEIG	735
RESULT 11			
ID	AAW51494		
XX	AAW51494	standard; protein; 735 AA.	
AC	AAW51494;		
XX	01-FEB-2002	(first entry)	
DE	Anthrax PA mutant K395D/K397D/D425K/D426K.		
XX			
KW	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;		
KW	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.		
XX			
OS	Bacillus anthracis.		
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference 395	/note= "Wild-type Lys substituted by Asp"	
FT	Misc-difference 397	/note= "Wild-type Lys substituted by Asp"	
FT	Misc-difference 425	/note= "Wild-type Asp substituted by Lys"	
FT	Misc-difference 426	/note= "Wild-type Asp substituted by Lys"	
XX			
PN	WO200182788-A2.		
XX			
PD	08-NOV-2001.		
XX			
PF	04-MAY-2001; 2001WO-US014372.		
XX			
PR	04-MAY-2000; 2000US-0201800P.		
XX			
PA	(HARD) HARVARD COLLEGE.		
XX			
PI	Collier RJ, Sellman BR;		
XX			
DR	WPI; 2002-017725/02.		
XX			
PT	Protecting humans against anthrax using mutant B groups (anthrax		
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus		
PT	anthracis.		
XX			
PS	Claim 4; Page; 77pp; English.		
XX			
CC	The invention relates to antibacterial agents comprising mutant forms of		
CC	pore-forming toxins (AAW52113 and AAW51484-AAW51500), especially mutants		
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B		
CC	moiety is anthrax protective antigen (PA) and using these mutants or		
CC	compositions of them for protecting against Bacillus anthracis infections		
CC	in humans, especially as vaccines. Note: The present sequence is not		
CC	given in the specification but is derived from the Bacillus anthracis		
CC	wild-type PA protein sequence shown in figure 13 (AAW51483)		
XX			
SQ	Sequence 735 AA;		
Query Match			
Best Local Similarity 99.7%; Score 3761; DB 5; Length 735;			
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF	60

Db	1	EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF	60
QY	61	QSAIWSGFTKYKKSDEYTFATSADNHVTWVDDQVINKASNSNKRLEKGRLYQIKQY	120
Db	61	QSAIWSGFTKYKKSDEYTFATSADNHVTWVDDQVINKASNSNKRLEKGRLYQIKQY	120
QY	121	QRENPTKGLDFKLYWTDSONKKEVSISSDNQLPELKQKSSNSRKRSTSAQTPVDRDN	180
Db	121	QRENPTKGLDFKLYWTDSONKKEVSISSDNQLPELKQKSSNSRKRSTSAQTPVDRDN	180
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSYDPEKVT	240
Db	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSYDPEKVT	240
QY	241	GRIDKNVSPGARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT	300
Db	241	GRIDKNVSPGARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT	300
QY	301	SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	301	SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRVYNTGTAPIYNNVLPPTSILVGLKNOTLATIKADENQLSQILAPNNYPSKNLAPIA	420
Db	361	NANIRVYNTGTAPIYNNVLPPTSILVGLKNOTLATIKADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKDFSPPTIMYNNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV	480
Db	421	LNAQKDFSPPTIMYNNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV	480
QY	481	LPOIQETTARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAGFNEPENGNL	540
Db	481	LPOIQETTARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAGFNEPENGNL	540
QY	541	QYQKCIDTEFDNFDDQTSQNIKNQAEALNATYVLDKIKLNAKONLILDKRPHYDR	600
Db	541	QYQKCIDTEFDNFDDQTSQNIKNQAEALNATYVLDKIKLNAKONLILDKRPHYDR	600
QY	601	NNIAGADESVKAEHREVNINSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY	660
Db	601	NNIAGADESVKAEHREVNINSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY	660
QY	661	DMLNISSLRODQKTFIDFKYNDKLPYISNPYKVNVAVTYKNTIINPSENGDTSTNG	720
Db	661	DMLNISSLRODQKTFIDFKYNDKLPYISNPYKVNVAVTYKNTIINPSENGDTSTNG	720
QY	721	IKKILIFSKKGYEIG	735
Db	721	IKKILIFSKKGYEIG	735
RESULT 12			
ID	AAW56959	standard; protein; 736 AA.	
XX			
AC	AAW56959;		
XX			
DT	25-APR-2000	(first entry)	
XX			
DE	B. anthracis MAT-PA protein.		
XX			
KW	Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;		
KW	tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.		
XX			
OS	Bacillus anthracis.		
XX			
PN	WO200002522-A2.		
XX			
PD	20-JAN-2000.		
XX			
PF	09-JUL-1999; 99WO-US015568.		
XX			
PR	10-JUL-1998; 98US-0092416P.		

XX (USMB-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;
XX WPI; 2000-182165/16.
XX N-PSDB; AA256875.
XX Recombinant DNA construct useful as vaccines for anthrax, in producing
XX host cells for analyzing the drugs and agents inhibiting anthrax.
XX Disclosure; Page 34; 35pp; English.
XX The invention provides a recombinant DNA construct that comprises a
XX vector and at least one nucleic acid (or its fragment) encoding a
XX combination of Bacillus anthracis proteins, selected from protective
XX antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA
XX with its secretory signals replaced with those of tissue plasminogen
XX activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine
XX for anthrax and in producing infectious alpha virus particles. These
XX particles, expressing the B. anthracis proteins are useful also as
XX vaccines for anthrax. Host cells transformed with the construct are
XX useful for analyzing the effectiveness of drugs and agents that inhibit
XX anthrax or B. anthracis proteins. The present sequence represents a B.
XX anthracis MAT-PA protein
XX Sequence 736 AA;
XX
XX Query Match 99.7%; Score 3761; DB 3; Length 736;
XX Best Local Similarity 99.7%; Pred. No. 3.1e-240;
XX Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
XX 2 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 61
XX 61 QSAIWSGFIKVKKSDYFATADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
XX 62 QSAIWSGFIKVKKSDYFATADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 121
XX 121 QRENTEKGLDFKLYWTDQNKKEVSSDNLQPELKQKSSNSRKRKSTASGTPVDRDN 180
XX 122 QRENTEKGLDFKLYWTDQNKKEVSSDNLQPELKQKSSNSRKRKSTASGTPVDRDN 181
XX 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPFKWSTASDPYDFEKT 240
XX 182 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPFKWSTASDPYDFEKT 241
XX 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNSTSRHT 300
XX 242 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNSTSRHT 301
XX 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
XX 302 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 361
XX 361 NANIRYVNTGTAPIYNNVPTSLVLGKNOTLATIKADENQLSQILAPNNYPSKNLPIA 420
XX 362 NANIRYVNTGTAPIYNNVPTSLVLGKNOTLATIKADENQLSQILAPNNYPSKNLPIA 421
XX 421 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNTFNGRVRVDTGNSWSEV 480
XX 422 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNTFNGRVRVDTGNSWSEV 481
XX 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSPLETTKPDMTLKEALKIATGFNEPKNL 540
XX 482 LPOIQETTARIIFNGKDLNLVERRIAANVPSPLETTKPDMTLKEALKIATGFNEPKNL 541
XX 541 QYGGKDIPTDFDQDTQNTKINQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600
XX 542 QYGGKDIPTDFDQDTQNTKINQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 601
XX 601 NNTAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660

Db 602 NNTAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 661
Qy 661 DMLNSSLRQDGKTFIDFKKYNIDKPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 662 DMLNSSLRQDGKTFIDFKKYNIDKPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 721
Qy 721 IKKILIESKGYEIG 735
Db 722 IKKILIESKGYEIG 736
RESULT 13
AAV56960
ID AAV56960 standard; protein; 763 AA.
XX AAV56960;
XX 25-APR-2000 (first entry)
DE B. anthracis TPA-PA protein.
XX Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;
XX tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.
XX Bacillus anthracis.
OS
XX WO200002522-A2.
XX 20-JAN-2000.
XX 09-JUL-1999; 99WO-US015568.
XX 10-JUL-1998; 98US-0092416P.
XX (USMB-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;
XX WPI; 2000-182165/16.
XX N-PSDB; AA256875.
XX Recombinant DNA construct useful as vaccines for anthrax, in producing
XX host cells for analyzing the drugs and agents inhibiting anthrax.
XX Disclosure; Page 32; 35pp; English.
XX The invention provides a recombinant DNA construct that comprises a
XX vector and at least one nucleic acid (or its fragment) encoding a
XX combination of Bacillus anthracis proteins, selected from protective
XX antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA
XX with its secretory signals replaced with those of tissue plasminogen
XX activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine
XX for anthrax and in producing infectious alpha virus particles. These
XX particles, expressing the B. anthracis proteins are useful also as
XX vaccines for anthrax. Host cells transformed with the construct are
XX useful for analyzing the effectiveness of drugs and agents that inhibit
XX anthrax or B. anthracis proteins. The present sequence represents a B.
XX anthracis TPA-PA protein
XX Sequence 763 AA;
XX
XX Query Match 99.7%; Score 3761; DB 3; Length 763;
XX Best Local Similarity 99.7%; Pred. No. 3.3e-240;
XX Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
XX 29 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 88
XX 61 QSAIWSGFIKVKKSDYFATADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
XX 89 QSAIWSGFIKVKKSDYFATADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 148

QY	121	QRENPTKGLDFKLYWTDSONKKEVVISDNLQLPKQKSSNRKRRKSTAGPTVPDRN	180
Db	149	QRENPTKGLDFKLYWTDSONKKEVVISDNLQLPKQKSSNRKRRKSTAGPTVPDRN	208
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT	240
Db	209	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT	268
QY	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT	300
Db	269	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT	328
QY	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTGLMTADTARL	360
Db	329	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTGLMTADTARL	388
QY	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA	420
Db	389	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA	448
QY	421	LNAQKDFSSPTITMNYNOFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGSNWSEV	480
Db	449	LNAQKDFSSPTITMNYNOFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGSNWSEV	508
QY	481	LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENGNL	540
Db	509	LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENGNL	568
QY	541	QYQKXDIETDFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPFYDR	600
Db	569	QYQKXDIETDFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPFYDR	628
QY	601	NNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIYVEIDTEGLKEVINDRY	660
Db	629	NNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIYVEIDTEGLKEVINDRY	688
QY	661	DLNLISSLRQDGKTIDFKKYNKDLPLYSNPNYKVNYYAVTKENTIINPSENGDTSTNG	720
Db	689	DLNLISSLRQDGKTIDFKKYNKDLPLYSNPNYKVNYYAVTKENTIINPSENGDTSTNG	748
QY	721	IKKILIFSKKGYEIG 735	
Db	749	IKKILIFSKKGYEIG 763	
RESULT 14			
ID	AAV56958	standard; protein; 764 AA.	
AC	AAV56958;		
DT	25-APR-2000	(first entry)	
XX	AAV56958;		
DE	B. anthracis protective antigen (PA) protein.		
XX	Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;		
KW	tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.		
XX	Bacillus anthracis.		
OS			
XX	WC0200002522-A2.		
PN			
XX	20-JAN-2000.		
PD			
XX	09-JUL-1999;	99WO-US015568.	
PF			
XX	10-JUL-1998;	98US-0092416P.	
PR			
XX	(USMB-) US MEDICAL RES INST INFECTIOUS DISEASES.		
PA			
XX	Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;		
PI			
XX			

DR	WPI; 2000-182165/16.	
DR	N-PSDB; AAZ56874.	
XX	Recombinant DNA construct useful as vaccines for anthrax, in producing	
PT	host cells for analyzing the drugs and agents inhibiting anthrax.	
XX	Disclosure; Page 33; 35pp; English.	
PS		
XX	The invention provides a recombinant DNA construct that comprises a	
CC	vector and at least one nucleic acid (or its fragment) encoding a	
CC	combination of Bacillus anthracis proteins, selected from protective	
CC	antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA	
CC	with its secretory signals replaced with those of tissue plasminogen	
CC	activator) and PA63 (63 Kda PA). The construct is useful as a DNA vaccine	
CC	for anthrax and in producing infectious alpha virus particles. These	
CC	particles, expressing the B. anthracis proteins are useful also as	
CC	vaccines for anthrax. Host cells transformed with the construct are	
CC	useful for analyzing the effectiveness of drugs and agents that inhibit	
CC	anthrax or B. anthracis proteins. The present sequence represents a B.	
CC	anthracis PA protein	
XX		
SQ	Sequence 764 AA;	
Query Match 99.7%; Score 3761; DB 3; Length 764;		
Best Local Similarity 99.7%; Pred. No. 3.3e-240;		
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY	1	EYKQENRLINESSESSQGLLYFSDLAFQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db	30	EYKQENRLINESSESSQGLLYFSDLAFQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 89
QY	61	QSAIWSGFTKVKKSDYEYTFATSDADNHVTMWVDDQEVINKASNNKIRLEKGRLYQIKY 120
Db	90	QSAIWSGFTKVKKSDYEYTFATSDADNHVTMWVDDQEVINKASNNKIRLEKGRLYQIKY 149
QY	121	QRENPTKGLDFKLYWTDSONKKEVVISDNLQLPKQKSSNRKRRKSTAGPTVPDRN 180
Db	150	QRENPTKGLDFKLYWTDSONKKEVVISDNLQLPKQKSSNRKRRKSTAGPTVPDRN 209
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240
Db	210	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT 269
QY	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Db	270	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 329
QY	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTGLMTADTARL 360
Db	330	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTGLMTADTARL 389
QY	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420
Db	390	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 449
QY	421	LNAQKDFSSPTITMNYNOFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGSNWSEV 480
Db	450	LNAQKDFSSPTITMNYNOFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGSNWSEV 509
QY	481	LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENGNL 540
Db	510	LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENGNL 569
QY	541	QYQKXDIETDFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPFYDR 600
Db	570	QYQKXDIETDFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPFYDR 629
QY	601	NNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIYVEIDTEGLKEVINDRY 660
Db	630	NNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYIYVEIDTEGLKEVINDRY 689
QY	661	DLNLISSLRQDGKTIDFKKYNKDLPLYSNPNYKVNYYAVTKENTIINPSENGDTSTNG 720

Db 690 DMLNISSLRQDGKTFIDFKYNDKPLVYISNPNYKVVAVTKENTIIINPSENGDTSTNG 749

QY 721 IKKILIFSKGYEIG 735

Db 750 IKKILIFSKGYEIG 764

RESULT 15

AA47306

ID AAB47306 standard; protein; 764 AA.

XX AAB47306;

AC 29-AUG-2001 (first entry)

XX Wild type B. anthracis protective antigen.

DE Lethal factor; LF; immunogen; LF4; protective antigen; PA; DNA vaccine;

XX humoral; cell-mediated; immune memory response.

XX Bacillus anthracis.

XX Key Location/Qualifiers

FT Peptide 1..29

FT /label= Signal peptide

FT /note= "Not given in the specification"

FT Protein 30..764

FT /label= PA

FT Peptide 204..764

FT /label= pCPA

XX WO200145639-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000WO-US034912.

XX 22-DEC-1999; 99US-01714599.

XX (OHIS) UNIV OHIO STATE RES FOUND.

PA (GALL/) GALLOWAY D R.

PA (MAYE/) MATECZUN A J.

XX Galloway DR, Mateczun AJ;

PI WPI: 2001-408540/43.

DR N-PSDB; AAC86016.

XX

PT Protecting animal against lethal infection with Bacillus anthracis, by

PT administering wildtype or mutated form of Bacillus anthracis lethal

PT factor protein or its fragment or a nucleic acid encoding the mutated

PT protein.

XX Claim 5; Fig 2; 33pp; English.

XX

CC This sequence shows the B. anthracis protective antigen (PA). An

CC immunogenic fragment of PA, pCPA, can be used to produce an immune

CC response which protects an animal against lethal infection with Bacillus

CC anthracis. DNA encoding the B. anthracis PA can be used in conjunction

CC with DNA encoding the lethal factor (LF) in a DNA vaccine. Using a DNA

CC vaccine which encodes the mutated LF protein or fragment alone or in

CC combination with a DNA encoding the PA protein or its fragment, both

CC components (humoral and cell-mediated) of the immune system are

CC stimulated, which results in longer term immune memory response. The

CC combined use of a mutated LF and PA gene or their fragments results in a

CC higher level of immune response, as judged by overall serum antibody

CC titers for LF and PA antigens, than the use of either LF or PA genes in

CC separate immunizations

XX

SQ Sequence 764 AA;

Query Match 99.7%; Score 3761; DB 4; Length 764;

Best Local Similarity 99.7%; Pred. No. 3.3e-240;

	Matches	733; Conservative	0; Mismatches	2; Indels	0; Gaps
QY	1	EVKQENRLINSESSSQGLGYFSDLNFOAPMVVTSTTGDLSIPSELENIPSENOYF	60		
Db	30	EVKQENRLINSESSSQGLGYFSDLNFOAPMVVTSTTGDLSIPSELENIPSENOYF	89		
QY	61	QSAIWSGFIKVKSDSEYTFATSAHNVMTWDDQEVINKASNNKIRLEKGRLYQIKIY	120		
Db	90	QSAIWSGFIKVKSDSEYTFATSAHNVMTWDDQEVINKASNNKIRLEKGRLYQIKIY	149		
QY	121	QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKRKSTASGPTVPDRDN	180		
Db	150	QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKRKSTASGPTVPDRDN	209		
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWSTASDPYDFEKT	240		
Db	210	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWSTASDPYDFEKT	269		
QY	241	GRIDKNVSPKARHPLVAAVPIVHVDNENIILSKNEDQSTQNTDSETRISKNTSRTHT	300		
Db	270	GRIDKNVSPKARHPLVAAVPIVHVDNENIILSKNEDQSTQNTDSETRISKNTSRTHT	329		
QY	301	SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL	360		
Db	330	SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL	389		
QY	361	NANIRYVNTGTAPIYNNVLTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA	420		
Db	390	NANIRYVNTGTAPIYNNVLTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA	449		
QY	421	LNAQKDFSTPTITMNYNQFLEKTKQLRLDQVYGNIAIYNFENGRVVDVTGNSWSEV	480		
Db	450	LNAQKDFSTPTITMNYNQFLEKTKQLRLDQVYGNIAIYNFENGRVVDVTGNSWSEV	509		
QY	481	LPOIQTETARIIFNGKOLNVERRIAANVPSPLETTKEDMTLKEALKIAGFNPNGL	540		
Db	510	LPOIQTETARIIFNGKOLNVERRIAANVPSPLETTKEDMTLKEALKIAGFNPNGL	569		
QY	541	QYQKDIETEFDFNFDOQTQSNIKNQLAELNATNIYVLDKIKLNAMNIIIRDKPFHYDR	600		
Db	570	QYQKDIETEFDFNFDOQTQSNIKNQLAELNATNIYVLDKIKLNAMNIIIRDKPFHYDR	629		
QY	601	NNIAGADESVVKEAHREVINSTEGLLLNIDKIRKILSGYVIEIETEGEKEVINDRY	660		
Db	630	NNIAGADESVVKEAHREVINSTEGLLLNIDKIRKILSGYVIEIETEGEKEVINDRY	689		
QY	661	DMLNIISSLRQDGKTFIDFKYNDKPLVYISNPNYKVVAVTKENTIIINPSENGDTSTNG	720		
Db	690	DMLNIISSLRQDGKTFIDFKYNDKPLVYISNPNYKVVAVTKENTIIINPSENGDTSTNG	749		
QY	721	IKKILIFSKGYEIG 735			
Db	750	IKKILIFSKGYEIG 764			

Search completed: May 3, 2004, 19:35:58

Job time : 52.1208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:31:43 ; Search time 15.0986 Seconds
(without alignments)
2513.152 Million cell updates/sec

Title: US-09-848-909A-10
Perfect score: 3774
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKILFFSKKGYEIG 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCFUS COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3761	99.7	735	1	US-08-021-601-4
2	3761	99.7	735	1	US-08-082-849B-4
3	3761	99.7	735	5	PCT-US94-01624-4
4	3716.5	98.5	903	1	US-08-021-601-12
5	3716.5	98.5	903	1	US-08-082-849B-12
6	3716.5	98.5	903	5	PCT-US94-01624-12
7	3584	95.0	719	1	US-08-082-849B-31
8	3584	95.0	719	5	PCT-US94-01624-31
9	1275	33.8	288	4	US-09-273-839A-8
10	775	20.5	881	3	US-08-960-780-32
11	775	20.5	881	3	US-09-073-898-32
12	775	20.5	881	4	US-09-307-106-8
13	775	20.5	881	4	US-09-850-351A-32
14	773	20.5	884	1	US-08-471-033-5
15	773	20.5	884	2	US-08-471-044-5
16	773	20.5	884	2	US-08-463-483A-5
17	773	20.5	884	2	US-08-471-046A-5
18	773	20.5	884	2	US-08-470-566B-5
19	773	20.5	884	2	US-08-469-334-5
20	773	20.5	884	3	US-09-300-529-5
21	773	20.5	1346	1	US-08-471-033-23
22	773	20.5	1346	2	US-08-471-044-23
23	773	20.5	1346	2	US-08-463-483A-23
24	773	20.5	1346	2	US-08-471-046A-23
25	773	20.5	1346	2	US-08-470-566B-23
26	773	20.5	1346	2	US-08-469-334-23
27	773	20.5	1346	3	US-09-300-529-23

28	772	20.5	852	1	US-08-471-033-36	Sequence 36, Appl
29	772	20.5	852	2	US-08-471-044-36	Sequence 36, Appl
30	772	20.5	852	2	US-08-463-483A-36	Sequence 36, Appl
31	772	20.5	852	2	US-08-471-046A-36	Sequence 36, Appl
32	772	20.5	852	2	US-08-470-566B-36	Sequence 36, Appl
33	772	20.5	852	2	US-08-469-334-36	Sequence 36, Appl
34	772	20.5	852	3	US-09-300-529-36	Sequence 36, Appl
35	772	20.5	1338	1	US-08-471-033-50	Sequence 50, Appl
36	772	20.5	1338	2	US-08-471-044-50	Sequence 50, Appl
37	772	20.5	1338	2	US-08-463-483A-50	Sequence 50, Appl
38	772	20.5	1338	2	US-08-471-046A-50	Sequence 50, Appl
39	772	20.5	1338	2	US-08-470-566B-50	Sequence 50, Appl
40	772	20.5	1338	2	US-08-469-334-50	Sequence 50, Appl
41	772	20.5	1338	3	US-09-300-529-50	Sequence 50, Appl
42	762.5	20.2	784	3	US-09-371-913A-7	Sequence 7, Appl
43	762.5	20.2	784	4	US-09-967-805-7	Sequence 48, Appl
44	753.5	20.0	860	4	US-09-307-106-48	Sequence 21, Appl
45	743.5	19.7	834	1	US-08-471-033-21	

ALIGNMENTS

RESULT 1
US-08-021-601-4
; Sequence 4, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 133 Carnegie Way, Suite 400
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021.601
; FILING DATE: 19930212
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-021-601-4

Query Match 99.7%; Score 3761; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 1.2e-254;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESQGLLGYFSDLNFPQMVVTSSTGDLSPSSSELNIPSENOYF 60
|||||

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDYFATSAADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYFATSAADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWDSQNKKEVSSDNIQLPELKQKSNKSKKSTAGPTVPDRN 180
Db 121 QRENTEKGLDFKLYWDSQNKKEVSSDNIQLPELKQKSNKSKKSTAGPTVPDRN 180
Qy 181 DGIPDSLEVGTVVVKKRTFLSPWISNIHEKKGLTKYKSPKSWSTASDPYDFEKVT 240
Db 181 DGIPDSLEVGTVVVKKRTFLSPWISNIHEKKGLTKYKSPKSWSTASDPYDFEKVT 240
Qy 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSTHT 300
Db 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSTHT 300
Qy 301 SEVHGNAEVAHAFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHAFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVFNENGRVVDTSNNSSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVFNENGRVVDTSNNSSEV 480
Qy 481 LPQIOETTARIIFNGKDLNVERRIAANPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540
Db 481 LPQIOETTARIIFNGKDLNVERRIAANPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540
Qy 541 QYQKDIITEFDNFDOOTSQNKQKLAELNATNIYVLDKIKNKAKNILLRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNKQKLAELNATNIYVLDKIKNKAKNILLRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTGLKEVINDRY 660

RESULT 2
US-082-849B-4
Sequence 4, Application US/08082849B
Patent No. 567274
GENERAL INFORMATION:
APPLICANT: Leppia, Stephen H.
APPLICANT: Klumpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
TITLE OF INVENTION: Related Methods
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-4
Query Match 99.7%; Score 3761; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 1.2e-254;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDYFATSAADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYFATSAADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWDSQNKKEVSSDNIQLPELKQKSNKSKKSTAGPTVPDRN 180
Db 121 QRENTEKGLDFKLYWDSQNKKEVSSDNIQLPELKQKSNKSKKSTAGPTVPDRN 180
Qy 181 DGIPDSLEVGTVVVKKRTFLSPWISNIHEKKGLTKYKSPKSWSTASDPYDFEKVT 240
Db 181 DGIPDSLEVGTVVVKKRTFLSPWISNIHEKKGLTKYKSPKSWSTASDPYDFEKVT 240
Qy 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSTHT 300
Db 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSTHT 300
Qy 301 SEVHGNAEVAHAFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHAFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVFNENGRVVDTSNNSSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVFNENGRVVDTSNNSSEV 480
Qy 481 LPQIOETTARIIFNGKDLNVERRIAANPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540
Db 481 LPQIOETTARIIFNGKDLNVERRIAANPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540
Qy 541 QYQKDIITEFDNFDOOTSQNKQKLAELNATNIYVLDKIKNKAKNILLRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNKQKLAELNATNIYVLDKIKNKAKNILLRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTGLKEVINDRY 660

QY 661 DMLNSSLRQDQKTFIDFKYNDKPLYSINPNYKVNYYAVTKENTINPSENGDTSTNG 720
DB 661 DMLNSSLRQDQKTFIDFKYNDKPLYSINPNYKVNYYAVTKENTINPSENGDTSTNG 720
QY 721 IKKILFSSKGYEIG 735
DB 721 IKKILFSSKGYEIG 735
RESULT 3
PCT-US94-01624-4
Sequence 4, Application PC/TUS9401624
GENERAL INFORMATION:
APPLICANT: Leppia, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
RELATED METHODS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSES: TOWNSEND AND TOWNSEND KOURIE and CREW
STREET: Steuart Street Tower, 20th Floor, One Market
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01624
FILING DATE: June 25, 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-4
Query Match 99.7%; Score 3761; DB 5; Length 735;
Best Local Similarity 99.7%; Pred. No. 1.2e-254;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVQENRLNSESQGLLYYFSDLNFPQPMVVTSTTGLSIPSSLENIPSENQYF 60
DB 1 EVQENRLNSESQGLLYYFSDLNFPQPMVVTSTTGLSIPSSLENIPSENQYF 60
QY 61 QSAIWGFIKVKSDVETATSDNHNVTWVDDQEVINKASNKIRLEKGRLYQIKQY 120
DB 61 QSAIWGFIKVKSDVETATSDNHNVTWVDDQEVINKASNKIRLEKGRLYQIKQY 120
QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNQLQPELKQSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKEVSSDNQLQPELKQSSNRKRSSTAGTVPDRDN 180
QY 181 DGIPTDSLEVEGYVDVKNKRTFSPWISNHEKKGITKYSPEKWTASDPSDFEYVT 240
DB 181 DGIPTDSLEVEGYVDVKNKRTFSPWISNHEKKGITKYSPEKWTASDPSDFEYVT 240

QY 241 GRIDKNVSEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTAE 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTAE 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDPSSPTITMNNYQFLEKTKQLRLDDVOYGNATYFNENGRVVRVDTGSNWSEV 480
DB 421 LNAQKDPSSPTITMNNYQFLEKTKQLRLDDVOYGNATYFNENGRVVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENG 540
QY 541 QYQKDIETBDFNFDQDQTSQNKQLAELNATNIYVLDKIKLNKQNLILRDKRFHYDR 600
DB 541 QYQKDIETBDFNFDQDQTSQNKQLAELNATNIYVLDKIKLNKQNLILRDKRFHYDR 600
QY 601 NNIAVGADESVMKAEHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKAEHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDQKTFIDFKYNDKPLYSINPNYKVNYYAVTKENTINPSENGDTSTNG 720
DB 661 DMLNSSLRQDQKTFIDFKYNDKPLYSINPNYKVNYYAVTKENTINPSENGDTSTNG 720
QY 721 IKKILFSSKGYEIG 735
DB 721 IKKILFSSKGYEIG 735
RESULT 4
US-08-021-601-12
Sequence 12, Application US/08021601
Patent No. 5591631
GENERAL INFORMATION:
APPLICANT: Leppia, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Nichols, Peter J.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
RELATED METHODS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSES: Needle & Rosenberg, P.C.
STREET: 133 Carnegie Way, Suite 400
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,601
FILING DATE: 19930212
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 903 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-021-601-12

Query Match 98.5%; Score 3716.5; DB 1; Length 903;
 Best Local Similarity 98.9%; Pred. No. 2e-251;
 Matches 726; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSLENIPISENQYF 60
 DB 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSLENIPISENQYF 60

QY 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
 DB 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLOLPKQKSSNSRKRSTAGPTVPDRN 180
 DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLOLPKQKSSNSRKRSTAGPTVPDRN 180

QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASDPYSDFEKVT 240
 DB 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASDPYSDFEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKTSTSRHT 300
 DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKTSTSRHT 300

QY 301 SEVHGNAEVHAFDFDGGVSAGFSNSSTVAIDHSLSLGERTWAEWTGLNTADTARL 360
 DB 301 SEVHGNAEVHAFDFDGGVSAGFSNSSTVAIDHSLSLGERTWAEWTGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
 DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITANNYQFLEKTKQLRLDTPQVIGNIATYFNFGREVDVTSNWEV 480
 DB 421 LNAQKDFSSPTITANNYQFLEKTKQLRLDTPQVIGNIATYFNFGREVDVTSNWEV 480

QY 481 LPQIQETTARIIPNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPQNL 540
 DB 481 LPQIQETTARIIPNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPQNL 540

QY 541 QYQKDIETDFDNFDQQTQNTKQNLAEINATNIYVLDKIKNAKMILIEDKPFHYDR 600
 DB 541 QYQKDIETDFDNFDQQTQNTKQNLAEINATNIYVLDKIKNAKMILIEDKPFHYDR 600

QY 601 NNIAVGADESVMKAEHREVINSSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
 DB 601 NNIAVGADESVMKAEHREVINSSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660

QY 661 DMLNISLRDQGTFFDFFKNDKLPYISNPNKYKNVAVTKENTIIINPSENGTSTNG 720
 DB 661 DMLNISLRDQGTFFDFFKNDKLPYISNPNKYKNVAVTKENTIIINPSENGTSTNG 720

QY 721 IKKIL---IFSKKG 731
 DB 721 IKKILKVVVLGKG 734

RESULT 5

US-08-082-849B-12
 Sequence 12, Application US/08082849B
 Patent No. 567274
 GENERAL INFORMATION:
 APPLICANT: Leppla, Stephen H.
 APPLICANT: Klimpel, Kurt R.
 APPLICANT: Arora, Naveen

APPLICANT: Singh, Yogendra
 APPLICANT: Nichols, Peter J.
 TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
 TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/082,849B
 FILING DATE: 25-JUN-1993
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/021,601
 FILING DATE: 12-FEB-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Weber, Kenneth A.
 REGISTRATION NUMBER: 31,677
 REFERENCE/DOCKET NUMBER: 15280-161-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 903 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-082-849B-12

Query Match 98.5%; Score 3716.5; DB 1; Length 903;
 Best Local Similarity 98.9%; Pred. No. 2e-251;
 Matches 726; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSLENIPISENQYF 60
 DB 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSLENIPISENQYF 60

QY 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
 DB 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLOLPKQKSSNSRKRSTAGPTVPDRN 180
 DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLOLPKQKSSNSRKRSTAGPTVPDRN 180

QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASDPYSDFEKVT 240
 DB 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASDPYSDFEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKTSTSRHT 300
 DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKTSTSRHT 300

QY 301 SEVHGNAEVHAFDFDGGVSAGFSNSSTVAIDHSLSLGERTWAEWTGLNTADTARL 360
 DB 301 SEVHGNAEVHAFDFDGGVSAGFSNSSTVAIDHSLSLGERTWAEWTGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
 DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITANNYQFLEKTKQLRLDTPQVIGNIATYFNFGREVDVTSNWEV 480

Db 421 LNAQDDFSTPTIMYNQFLELEKTKQLRLDTQVYGNIAYNFENGVRVDTGSNNSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVLVERRIAAVNPSDPLETTKPDMTLKEALKIARFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVLVERRIAAVNPSDPLETTKPDMTLKEALKIARFNEPNGNL 540
Qy 541 QYQGDITFEFDFNFQOQTSQNIKNQLAELNATNIYVLDKIKNAKWNILIRDKGFHYDR 600
Db 541 QYQGDITFEFDFNFQOQTSQNIKNQLAELNATNIYVLDKIKNAKWNILIRDKGFHYDR 600
Qy 601 NNIAGADESVVKEAHRVINSSTEGLLNIDKQIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHRVINSSTEGLLNIDKQIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNKYVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNKYVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKIL---IFSKKG 731
Db 721 IKKILKKVVLGKG 734

RESULT 6
PCT-US94-01624-12
; Sequence 12, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Lepple, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE and CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; Plaza
; City: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; Filing Date: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01624-12

Query Match 98.5%; Score 3716.5; DB 5; Length 903;
Best Local Similarity 98.9%; Pred. No. 2e-251;
Matches 726; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

Qy 1 EVKQENLLNESESSQGLLYYFSDLNFPQPMVVTSTTGLSLPSSSELENIIPSENQYF 60
61 QSAIMSGFIKVKKSDSEYTPATSAADNHVTMMVDDQEVINKASNSNKIRLEKGLYQIKQY 120
61 QSAIMSGFIKVKKSDSEYTPATSAADNHVTMMVDDQEVINKASNSNKIRLEKGLYQIKQY 120
121 QRENPTKEGLDFKLYWTDSONKKEVISSDNLOIPELKQKSSNSRKKRSTASAGTVPDRN 180
121 QRENPTKEGLDFKLYWTDSONKKEVISSDNLOIPELKQKSSNSRKKRSTASAGTVPDRN 180
181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKKGLTKYKSSPEKMWSTASDPYSPEKVT 240
181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKKGLTKYKSSPEKMWSTASDPYSPEKVT 240
241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTTSKNTSRTH 300
241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTTSKNTSRTH 300
301 SEVHGNAEVAHFFDIDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
301 SEVHGNAEVAHFFDIDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420
361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420
421 LNAQDDFSTPTIMYNQFLELEKTKQLRLDTQVYGNIAYNFENGVRVDTGSNNSEV 480
421 LNAQDDFSTPTIMYNQFLELEKTKQLRLDTQVYGNIAYNFENGVRVDTGSNNSEV 480
481 LPOIQTETARIIFNGKDLNVLVERRIAAVNPSDPLETTKPDMTLKEALKIARFNEPNGNL 540
481 LPOIQTETARIIFNGKDLNVLVERRIAAVNPSDPLETTKPDMTLKEALKIARFNEPNGNL 540
541 QYQGDITFEFDFNFQOQTSQNIKNQLAELNATNIYVLDKIKNAKWNILIRDKGFHYDR 600
541 QYQGDITFEFDFNFQOQTSQNIKNQLAELNATNIYVLDKIKNAKWNILIRDKGFHYDR 600
601 NNIAGADESVVKEAHRVINSSTEGLLNIDKQIRKILSGYVIEIDTEGLKEVINDRY 660
601 NNIAGADESVVKEAHRVINSSTEGLLNIDKQIRKILSGYVIEIDTEGLKEVINDRY 660
661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNKYVAVTKENTIINPSENGDTSTNG 720
661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNKYVAVTKENTIINPSENGDTSTNG 720
721 IKKIL---IFSKKG 731
721 IKKILKKVVLGKG 734

RESULT 7
US-08-082-849B-31
; Sequence 31, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Lepple, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; City: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/082,849B
FILING DATE: 25-JUN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,601
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-161-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-082-849B-31

Query Match 95.0%; Score 3584; DB 1; Length 719;
Best Local Similarity 95.3%; Pred. No. 2,6e-242;
Matches 704; Conservative 4; Mismatches 7; Indels 24; Gaps 2;
QY 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDYVTSADNHVTVMVDDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYVTSADNHVTVMVDDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNS----RKGRSTSAGPTVP 176
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSATIMMQRGFLQGPVTP 180
QY 177 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYSD 236
DB 181 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYSD 240
QY 237 EKYTGRIKQVSEARHPLVAAPYIVHVDYENIILSKNEDQSTQNTDSETRISKNWTS 296
DB 241 EKYTGRIKQVSEARHPLVAAPYIVHVDYENIILSKNEDQSTQNTDSETRISKNWTS 300
QY 297 RTHTSEVHGNAEVHAFDIDGGSVAGFNSNSSTVAIDHSLSLAGERTWAETMGLTAD 356
DB 301 RTHTSEVHGNAEVHAFDIDGGSVAGFNSNSSTVAIDHSLSLAGERTWAETMGLTAD 360
QY 357 TARIANIRVYNTGTAPIYVNLPTTSILGKQNTLAKADENQSLQILAPNNYPSKUL 416
DB 361 TARIANIRVYNTGTAPIYVNLPTTSILGKQNTLAKADENQSLQILAPNNYPSKUL 420
QY 417 APIALNAQDFFSPTITMN-----YGNIAIYFNFGVRVDTGSN 476
DB 421 APIALNAQDFFSPTITMN-----YGNIAIYFNFGVRVDTGSN 460
QY 477 WSEVLPOIQTETARIIFNGKDLNLAVERRIAANVPSDPLETTKEDMTLKEALKIATGENEP 536
DB 461 WSEVLPOIQTETARIIFNGKDLNLAVERRIAANVPSDPLETTKEDMTLKEALKIATGENEP 520
QY 537 NGNLQYQKDIETBFDFNFDOQTSONIKNQLAEINATNIYTVLDKIKLNAMNILLRDKRF 596
DB 521 NGNLQYQKDIETBFDFNFDOQTSONIKNQLAEINATNIYTVLDKIKLNAMNILLRDKRF 580
QY 597 HYDRNNIACADESVYKAEHREVINSSTGLIILNIDKIRKILSGYVIEIETEGLEKEVI 556
DB 581 HYDRNNIACADESVYKAEHREVINSSTGLIILNIDKIRKILSGYVIEIETEGLEKEVI 640

QY 657 NDRYDMLNLSLRQDGKTFIDFKKYNKDKLPYLSHPNPKYKVVAVTKENTINPSENGDT 716
DB 641 NDRYDMLNLSLRQDGKTFIDFKKYNKDKLPYLSHPNPKYKVVAVTKENTINPSENGDT 700
QY 717 STNGIKKILIFSKKGYEIG 735
DB 701 STNGIKKILIFSKKGYEIG 719
RESULT 8
PCT-US94-01624-31
Sequence 31: Application PC/TUS9401624
GENERAL INFORMATION:
APPLICANT: Leppla, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Arora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
TITLE OF INVENTION: RELATED METHODS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
STREET: Stewart Street Tower, 20th Floor, One Market
STREET: Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01624
FILING DATE: June 25, 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-31

Query Match 95.0%; Score 3584; DB 5; Length 719;
Best Local Similarity 95.3%; Pred. No. 2,6e-242;
Matches 704; Conservative 4; Mismatches 7; Indels 24; Gaps 2;
QY 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDYVTSADNHVTVMVDDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYVTSADNHVTVMVDDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNS----RKGRSTSAGPTVP 176
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSATIMMQRGFLQGPVTP 180
QY 177 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYSD 236
DB 181 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYSD 240

QY 237 EKVTRIDKXVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTSKNTSTS 296
DB 241 EKVTRIDKXVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTSKNTSTS 300
QY 297 RTHSEVHGNAEVHASFDDIGSVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTAD 356
DB 301 RTHSEVHGNAEVHASFDDIGSVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTAD 360
QY 357 TARLNANIRYNTGTAPIYVLPITSLVLGKNOQLATIKADENOLSOILAPNNYPPSKNL 416
DB 361 TARLNANIRYNTGTAPIYVLPITSLVLGKNOQLATIKAKENOLSOILAPNNYPPSKNL 420
QY 417 APIALNAQKQSSPTITMNNYQFLEKTKQLRLDTPQVGNIAATYFNGRVRVDTGSN 476
DB 421 APIALNAQKQSSPTITMNNYQFLEKTKQLRLDTPQVGNIAATYFNGRVRVDTGSN 460
QY 477 WSEVLPOIQTETARTIIFNGKDLNVERRIAAVNSPDPLETTKPDMTLKEALKIAGFNEP 536
DB 461 WSEVLPOIQTETARTIIFNGKDLNVERRIAAVNSPDPLETTKPDMTLKEALKIAGFNEP 520
QY 537 NGNLQYQKDIETFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKXNNILIRDKRF 596
DB 521 NGNLQYQKDIETFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKXNNILIRDKRF 580
QY 597 HYDRNNAIVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVI 656
DB 581 HYDRNNAIVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVI 640
QY 657 NDYRQDMLNLSLRQDGKTFIDFKKYNKDLPLYISPNYKVNVAVTKNTIINPSENGDT 716
DB 641 NDYRQDMLNLSLRQDGKTFIDFKKYNKDLPLYISPNYKVNVAVTKNTIINPSENGDT 700
QY 717 STNGIKKILIFSCKGYEIG 735
DB 701 STNGIKKILIFSCKGYEIG 719

RESULT 9
US-09-273-839A-8
; Sequence 8, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-273-839A-8

Query Match 33.8%; Score 1275; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 1.7e-81; Mismatches 0; Indels 0; Gaps 0;
Matches 250; Conservative 0

QY 486 ETTARIIFNGKDLNVERRIAAVNSPDPLETTKPDMTLKEALKIAGFNEPNGNLQYQK 545
DB 29 ETTARIIFNGKDLNVERRIAAVNSPDPLETTKPDMTLKEALKIAGFNEPNGNLQYQK 88
QY 546 DITEFPNFDQTSQNIKNQLAELNATNIYTVLDKIKLNKXNNILIRDKRFHYDRNTAV 605
DB 89 DITEFPNFDQTSQNIKNQLAELNATNIYTVLDKIKLNKXNNILIRDKRFHYDRNTAV 148
QY 606 GADSVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVINDRYDMLNI 665
DB 149 GADSVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVINDRYDMLNI 208

QY 666 SSLRODGKTFIDFKKYNKDLPLYISPNYKVNVAVTKNTIINPSENGDTSTNGIKKIL 725
DB 209 SSLRODGKTFIDFKKYNKDLPLYISPNYKVNVAVTKNTIINPSENGDTSTNGIKKIL 268
QY 726 IFSKKGYEIG 735
DB 269 IFSKKGYEIG 278

RESULT 10
US-08-960-780-32
; Sequence 32, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Gerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmelts, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA-708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 177C8
; US-08-960-780-32

Query Match 20.5%; Score 775; DB 3; Length 881;
Best Local Similarity 30.2%; Pred. No. 7.2e-46;
Matches 244; Conservative 134; Mismatches 299; Indels 130; Gaps 34;
QY 4 QENRLLNESSSQGLLGYYFSDLNFCAPMVVTSITGDLISPSSELEN--IPSENCYFQ 61
DB 39 QRNQ--QKMDRKGLLGYYFKGKDF-SNLTMPFTRDSTLIYDQQTANKLLKKQOEYQ 94
QY 62 SAINSGFIKVKSDVETFATSDNHTVMWDDQEVINKASNNKIRLEKRLYQIKIQYQ 121

Db 95 SIRWIGLIQSKETGDFTNLSEDEQAIEINGKIISNKGKEQVHVLEKGLVPKIEYQ 154
Qy 122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNLQLPKLOKSS-----N 162
Db 155 SD--TKFNIDSKTPEKELFKIDSONQPOQVQOQDELNPFNKESQEFLLAKSKINLFT 212
Qy 163 SRKRSTASGPTVDRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSS 222
Db 213 QKMKREIDED--TDTGDSIPDLWEENGYTI-----QNRIAVKWDDSL-ASKGYTKFVSN 264
Qy 223 PEKSTASDPYSDPEKVTGRDKNVSPHARPLVAAPVIVHVDMENTIILSKNEDQSTQNT 282
Db 265 PLESHTVGDPYDYKAAARDLJLSNAKETFNPLVAAPFVSVNMEKVLSPNENLS----- 320
Qy 283 DSETRTISKNTSTSRHTSEVHGNAEVSASFDDIGGSVSAGFSNSNS--TVAIDHSL 339
Db 321 ---NSVESHSSTNWSYI-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 363
Qy 340 LAGERTWAETWG-----LNTADTARLANIRVYVNTGTAPIYVNLPTTSLVLGKNOTLATIK 395
Db 364 VAQE--WGTSGNTSQFNTASAGYLNANVRVNYVGTGAIVDVKPTTSFVL--NNDTIAIT 420
Qy 396 ADENQLSOILAPNNYPSKLNAPALNAQKDFSSPTITMNYNQFLEKTKQLRLDQV 455
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHSPITLKKQVDNLLNKKPMWLETNQT 480
Qy 456 YGNATYFNGRVRVDTGWSWSEVLPQIQTTHARILFNGKDLNLFVRRVAAVNPSPLE 515
Db 481 DG---VYKIKDTHGNIVTGGWNGVIOQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPED 536
Qy 516 TTKPDMTLKEALKIAF--GFNEPNGNLOVQKQDTEFD--NFDQOTSQNIKNQLABL-- 569
Db 537 KT-PSLTLDKALKYPPDEIEGLLYKKNPIYESSVMYLDENTAKETVKQLNDTGT 595
Qy 570 ---NATNIYVUDKIKLNAKONILIRDKRFHYDRNNAIAGDAESVVKHAEVINSSTEG 626
Db 596 KFKDVSHLYDV---KLTPQNVITIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647
Qy 627 -----LLNID-----KDIRKILSGYIVEIEDTE-----GLKE 654
Db 648 KKOYSSNPNPALNTLTDQAEKLNORDYISLYMKSEKNTQCEITIDGEIYPTTKTVN 707
Qy 655 VINDRYDMLNI--SSLROGKTFIDFKYNDKPLIYINPNVKNVYVNTKENTINPSE 712
Db 708 VNKNYKRLDIIAHNIKSPISSIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSE 763
Qy 713 NGDT-STNGIK---KLIFSKKGYEIG 735
Db 764 IKQIVSRYGKLEDDGLIDKKGHIHV 790

RESULT 11
US-09-073-898-32
Sequence 32, Application US/09073898
Patent No. 6242669
GENERAL INFORMATION:
APPLICANT: Feitelson, Jerald S.
APPLICANT: Schrepp, H. Ernest
APPLICANT: Narva, Kenneth E.
APPLICANT: Stockhoff, Brian A.
APPLICANT: Schmeits, James
APPLICANT: Loewer, David
APPLICANT: Dullum, Charles Joseph
APPLICANT: Muller-Cohn, Judy
APPLICANT: Stamp, Lisa
APPLICANT: Morrill, George
APPLICANT: Finstad-Lee, Stacey
TITLE OF INVENTION: No. 6242669el Pesticidal Toxins and Nucleotide
TITLE OF INVENTION: Sequences Which Encode These Toxins
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606-6669
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,898
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,848
FILING DATE: 30-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/960,780
FILING DATE: 30-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-708C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 881 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: PS177C8
US-09-073-898-32

Query Match 20.5%; Score 775; DB 3; Length 881;
Best Local Similarity 30.2%; Pred. No. 7.2e-46;
Matches 244; Conservative 134; Mismatches 299; Indels 130; Gaps 34;

Qy 4 QENRLNSESQGLLGYFSDLNFAQPMWVTSSTTGDLSIPSSSELEN--IPSENQYEQ 61
Db 39 QKQK---QKMDRKGGLGYFKGKDF-SNLTWFAPTRDTLLYDQQTANKLLDKQEQVQ 94
Qy 62 SAWSGFIYKKSDEYTFATSNADNHVVMVDQEVINKASNKIRLEKRLYQIKIQV 121
Db 95 SIRWIGLIQSKETGDFTNLSEDEQAIEINGKIISNKGKEQVHVLEKGLVPKIEYQ 154
Qy 122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNLQLPKLOKSS-----N 162
Db 155 SD--TKFNIDSKTPEKELFKIDSONQPOQVQOQDELNPFNKESQEFLLAKSKINLFT 212
Qy 163 SRKRSTASGPTVDRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSS 222
Db 213 QKMKREIDED--TDTGDSIPDLWEENGYTI-----QNRIAVKWDDSL-ASKGYTKFVSN 264
Qy 223 PEKSTASDPYSDPEKVTGRDKNVSPHARPLVAAPVIVHVDMENTIILSKNEDQSTQNT 282
Db 265 PLESHTVGDPYDYKAAARDLJLSNAKETFNPLVAAPFVSVNMEKVLSPNENLS----- 320
Qy 283 DSETRTISKNTSTSRHTSEVHGNAEVSASFDDIGGSVSAGFSNSNS--TVAIDHSL 339
Db 321 ---NSVESHSSTNWSYI-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 363
Qy 340 LAGERTWAETWG-----LNTADTARLANIRVYVNTGTAPIYVNLPTTSLVLGKNOTLATIK 395
Db 364 VAQE--WGTSGNTSQFNTASAGYLNANVRVNYVGTGAIVDVKPTTSFVL--NNDTIAIT 420
Qy 396 ADENQLSOILAPNNYPSKLNAPALNAQKDFSSPTITMNYNQFLEKTKQLRLDQV 455
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHSPITLKKQVDNLLNKKPMWLETNQT 480

Db 764 IKQYSGYKLEBGLIDKGGIHYG 790

RESULT 13

US-09-850-351A-32

Sequence 32, Application US/09850351A

Patent No. 6656908

GENERAL INFORMATION:

APPLICANT: Feltelson, Jerald S.

Schnepp, H. Ernest

Narva, Kenneth E.

Stockhoff, Brian A.

Schmeits, James

Loewer, David

Dullum, Charles Joseph

Muller-cohn, Judy

Stamp, Lisa

Morrill, George

TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide Sequences Which Encode These Toxins

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606-6669

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/850,351A

FILING DATE: 07-May-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/073,898

FILING DATE: 06-MAY-1998

APPLICATION NUMBER: US 08/960,780

FILING DATE: 30-OCT-1997

APPLICATION NUMBER: US 60/029,848

FILING DATE: 30-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sanders, Jay M.

REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708CD1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 881 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: PS177C8

SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-09-850-351A-32

Query Match 20.5%; Score 775; DB 4; Length 881;

Best Local Similarity 30.2%; Pred. No. 7.2e-46;

Matches 244; Conservative 13; Mismatches 299; Indels 130; Gaps 34;

Qy 4 QENRLNESSSQGLGYFSDINFOAPMVVTSSTGDSIPSSLEN--IPSENQYFQ 61

Db 39 QKQK--QKEMDRKGLGYFKGKDF-SNLTMPAPTRDSPLIYDQQTANKLIDKQKQYQ 94

Qy 62 SATWSGFIKYSDEYFATSAADNHTVMVDDQEVINKASNSKIRLEKRLVQIKIQY 121

Db 95 SIRWIGLIQSKETGDFTFNLSEDEQAITEINGKIISNKGKQVHVHLEKGLVPIKIEYQ 154

Qy 122 RENPTEKGLD-----FKLYWTDSONKEVSSNQLQPLKQKSS-----N 162

Db 155 SD--TKFNIDSKTFKELKLFIDSONQFQQVQDELNRPENFKKESQEFLLAKPSKINLFT 212

Qy 163 SRKKESTAGPTVPDRDNDGIPDSLEVEGYTVDVKNKTFILSPWISNHEKKGITKYSKS 222

Db 213 QRMKEIDED--TDTDGSIPDLWEENGYTI---QNRIAVKWDDSL-ASKGYTKFVSN 264

Qy 223 PEKWTASDPYSDPEKVTGRIDKNVSPARHPLVAAPVIVHVDMMENIILSKNEDOSTQNT 282

Db 265 PLESHTVGDPTDYDEKAARDLDSNAKETFNPLVAAPFVSNVSMKVLSPNENUS----- 320

Qy 283 DSETRTISKNTSTSTRTHTSEVHGNAEIVHASFPDICGVSAGFSNSNS--TVAIDHLS 339

Db 321 ----NSVESHSSTNWSYT-----NTE-----CASVEAGICPGKISFGVSVNYQHSST 363

Qy 340 LAGERTWARTMG----LNTADTARLANANIRYNTGTAPVNVLPITTSVLGKNQTLATIK 395

Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANRYNVNVTGAIYDVKPTISFVL--NNDIATIT 420

Qy 396 ADENQLSQILAPNNYPSKNLAPIALNAQKPSSTPITMNNYQFLELEKTKQLRLDQV 455

Db 421 AKSNSTALNISGESYPPKGGQGLAITSDDFNSHPITLKKQVDNLLNKKPMLETNOT 480

Qy 456 YGNIATYFNGRVRVDTGNSWSEVLPOIQTETARIIFNGKDLNLVERIAAVNPSDPLE 515

Db 481 DG---VYKIKDTHGNIYVTCGEWNGVIQIKAKTASIIVDDGE-RVAEKRVAAKDYENPED 536

Qy 516 TTKPDMTLKALKIAP--GFNEPNGNLOYQGDITFDF--NFDQOTSONIKNQLAEL-- 569

Db 537 KT-PSUTLKDALKLSYDPDEIKIEGLLYYKNPIYESSVMTYLDENTAKEVTKQLNDTTG 595

Qy 570 ---NATNIYTVLDKIKLNAMNIIILDRKRFHVORNNIAGADESVVKEAHRVINSSTEG 626

Db 596 KFKDVSHLYDV---KLTPEKMNVTIK-LSILYDN---AESNDNSIGKNTNTNIVSGNGNG 647

Qy 627 -----LLNLID-----KDIRKILSYIVEIEDTE-----GLKE 654

Db 648 KKOYSSNPNPANLTLNTDAQEKLNRDYYISLYMKSEKNTQCEITIDGEIYPIITTKYN 707

Qy 655 VINDRYDMLNI--SSLRQDGKTFIDPKYNDKPLVYISNPNYKVNKYATVKENTINPSE 712

Db 708 VNKNYKRLDIIAHTNKSNIPISSHIKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763

Qy 713 NGDT-STNGIK---KILIFSKKGYEIG 735

Db 764 IKQYSGYKLEBGLIDKGGIHYG 790

RESULT 14

US-08-471-033-5

Sequence 5, Application US/08471033

Patent No. 5770896

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W

APPLICANT: Koziel, Michael G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J

APPLICANT: Carr, Brian

APPLICANT: Dessai, Nalini M

APPLICANT: Kostichka, N. Kristy

APPLICANT: Duck, Nicholas B

APPLICANT: Estruch, Juan J

TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 884 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-044-5

Query Match 20.5%; Score 773; DB 2; Length 884;
Best Local Similarity 30.2%; Pred.No.1e-45;
Matches 246; Conservative 129; Mismatches 294; Indels 146; Gaps 34;

QY	4	QENRLNESSESSQGLGYFFDLNFCAPMVVTSSTGDLSPSSELEN--IPSENQYFQ	61
Db	42	QKQKQ--QKEMDRKGLGYFFKGF-SNLTMPAPTRDSTLIYDQQTANKLLDKKQEQY	97
QY	62	SAIWSGFIKVKSDGYEATSDNHYTMVDDQEVINKASNSNKIRLEKGRLYQIKIYO	121
Db	98	STRWIGLISGKETGDTFFNLSEDEQAIIEINGKILSNKGEKQVHVHLEKGLVPIKIBYQ	157
QY	122	RENPTKGLD-----FKLYWTDSONKKEVISSDNLQLPELKQKSS-----N	162
Db	158	SD--TKFNI DSKTFKELKLFKIDSQNPQVQVQDELARNPEFNKESQEF LAKPSKINLFT	215
QY	163	SRKRSTSGAPVDPDRNDGIDPSLVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSS	222
Db	216	QMKREIDED---TDTGDSIPDLWEENGYTI-----QNRIVAKWDDSL-ASKGYTKFVSN	267
QY	223	PEKWTASDPYSDFEKTGRIDKNVSPPEARHPLVAAPIYVHVDMENIILSKNEDOSTQNT	282
Db	268	PLESHTVGPDPYDYEKARDLDSNAKETFNPLVAAFPSVNVSMKVILSPNENLS----	323
QY	283	DSETRISKNSTSRTHTEVHGNAEVHASFDDIGGSVAGSPNSNSS---TVADHSL	339
Db	324	---NSVESHSSTNWSYT-----NTE-----CASVEAGIGPKGISFGVSVNYQHS	366
QY	340	LAGERTWAETMG---LNTADTARLNANIRYVNTGPARIYVLPPTSGLGNQOTLATIK	395
Db	367	VAGE--WGSTGNSTQNTASAGYLNANRYNVNVTGAIYDVKPTSPVL--NDTIIAIT	423
QY	396	ADENQLSQILAPNRYPSKNLAPIALNAQKDFSSPTITMNVNQFLELEKTKQLRLDTQV	455
Db	424	AKSNSTALNISPGESYFKKGONGIAITSMDDFNHPITLKKQVDNLLNKKPMLETNQ	483
QY	456	YGNIAVYFNENGRVRYDTGNSMSEVLPOIQTETARIIFNGKDLNLVERRIAAVNP	515
Db	484	DG---VYKIKDTHGNIVTGEWNGVIQQIKAKTASIIVDGGE-RVAEKRVAAKDYEN	539
QY	516	TKPDMTLKALKIAP--GFNEPNGNLQYQKDIETEDF--NFDQQTSONIKNQLAEL--	569
Db	540	KT--PSLTLDALKLSPDEIKETEGLLYKKNPIYESSVMTYLDENTAKEVTKQLNDTTG	598
QY	570	---NATNIYTVLDKIKLNAKMLIRDKPFHYDRNNIAGADESVVKEAHREVINS	626
Db	599	KFDVSHLYDV-----KLTFRMNVTK-LSILYDN---AESDNSIGKWTNTNIVSGG	650
QY	627	-----LLNID-----KDIRKILSGYIVEIDTE-----GLKE	654
Db	651	KKOYSNNPDANLTNTDAQEKLNKNRDYIISLYMKSEKNTQCEITDGEIPIITKT	710
QY	655	VINDRYDMLN-----ISSLRQDGKTFIDFKKYNKLELYISNPNYKNVYATVKE	704
Db	711	VNKDNYKRLDITAHNKSPISSLH-----IKTNDDEITLFWDDISI-TDVASIK	760
QY	705	NTIINPSENGDT--STNGIK---KILIFSKGYEIG	735
Db	761	N--LTDSEIKQIYSRYGKLEDCILLDKKGIHYG	793

Search completed: May 3, 2004, 19:42:42
Job time : 17.0986 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:40:13 ; Search time 36.6393 Seconds
(without alignments)
5560.545 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENRLNESSSQGLL.....TSTNGIKKILIFSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277199581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3761	99.7	735	15	US-10-410-647-30
2	3761	99.7	735	12	US-09-848-909-1
3	3761	99.7	735	12	US-09-848-909-2
4	3761	99.7	735	12	US-09-848-909-3
5	3761	99.7	735	12	US-09-848-909-4
6	3761	99.7	735	12	US-09-848-909-5
7	3761	99.7	735	12	US-09-848-909-6
8	3761	99.7	735	12	US-09-848-909-7
9	3761	99.7	735	12	US-09-848-909-8
10	3761	99.7	735	12	US-09-848-909-9
11	3761	99.7	735	12	US-09-848-909-10
12	3761	99.7	735	12	US-09-848-909-11
13	3761	99.7	735	12	US-09-848-909-12
14	3761	99.7	735	12	US-09-848-909-13
15	3761	99.7	735	12	US-09-848-909-14

16	3761	99.7	736	12	US-09-848-909-15	Sequence 15, Appl
17	3761	99.7	736	12	US-09-848-909-16	Sequence 16, Appl
18	3761	99.7	736	12	US-09-848-909-17	Sequence 17, Appl
19	3761	99.7	736	12	US-09-848-909-18	Sequence 18, Appl
20	3761	99.7	736	12	US-09-848-909-19	Sequence 19, Appl
21	3761	99.7	736	12	US-09-848-909-20	Sequence 20, Appl
22	3761	99.7	736	12	US-09-848-909-21	Sequence 21, Appl
23	3761	99.7	736	12	US-09-848-909-22	Sequence 22, Appl
24	3761	99.7	736	15	US-10-442-502-7	Sequence 5, Appl
25	3761	99.7	763	15	US-10-442-502-5	Sequence 7, Appl
26	3761	99.7	763	15	US-10-442-502-6	Sequence 6, Appl
27	3758	99.6	735	12	US-10-402-466A-9	Sequence 9, Appl
28	3758	99.6	735	12	US-10-402-466A-13	Sequence 13, Appl
29	3758	99.6	735	14	US-10-332-282-13	Sequence 13, Appl
30	3754	99.5	735	12	US-09-848-909-30	Sequence 30, Appl
31	3754	99.5	764	12	US-10-253-286-681	Sequence 681, App
32	3754	99.5	764	15	US-10-245-871-681	Sequence 681, App
33	3654	96.8	764	9	US-09-747-521-4	Sequence 4, Appl
34	3654	96.8	764	13	US-10-106-014-4	Sequence 4, Appl
35	3654	96.8	764	13	US-10-105-695-4	Sequence 4, Appl
36	3654	96.8	764	14	US-10-105-694-4	Sequence 4, Appl
37	3050	80.8	599	12	US-09-848-909-24	Sequence 24, Appl
38	3045	80.7	595	14	US-10-332-282-11	Sequence 11, Appl
39	2918	77.3	573	12	US-10-402-466A-22	Sequence 22, Appl
40	2906	77.0	569	15	US-10-442-502-8	Sequence 8, Appl
41	2496	66.1	487	14	US-10-332-282-5	Sequence 5, Appl
42	2181	57.8	426	14	US-10-332-282-9	Sequence 9, Appl
43	2135	56.6	423	12	US-10-402-466A-24	Sequence 24, Appl
44	1632	43.2	318	14	US-10-332-282-7	Sequence 7, Appl
45	1338	35.5	258	14	US-10-332-282-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-410-647-30
; Sequence 30, Application US/10410647
; Publication No. US20030235818A1
; GENERAL INFORMATION:
; APPLICANT: PLEXUS VACCINE, INC.
; APPLICANT: Katritsch, Vsevolod
; APPLICANT: Bordner, Andrew
; APPLICANT: Deans, Robert
; APPLICANT: Sumner, Mary
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES, AND METHOD OF IDENTIFYING SAME
; FILE REFERENCE: PLEX1110-1
; CURRENT APPLICATION NUMBER: US/10/410,647
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/373,668
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/371,256
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/371,250
; PRIOR FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 30
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-410-647-30

Query Match 99.7%; Score 3761; DB 15; Length 735;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLLYGFSDLNFAQPMVVTSSTTGDLSTPSSSELENIPISENQVF 60

Db 1 EVKQENRLNESSSQGLLYGFSDLNFAQPMVVTSSTTGDLSTPSSSELENIPISENQIF 60

QY 61 QSAIWSGFIVKKKSDSYTFATSDADNHVTWVDQEVINKASNKIRLEKGLYQIKQY 120

Db 61 QSAIWSGFIKVKSDBYEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNQLQPLKQKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNQLQPLKQKSSNSRKRSTSGPTVPDRDN 180
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSSTASDYSDFEKT 240
Db 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSSTASDYSDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHAGFFDIGSVSAGFSNSSTVAIDHSLSLAGERTWAGTGLTADTARL 360
Db 301 SEVHGNAEVHAGFFDIGSVSAGFSNSSTVAIDHSLSLAGERTWAGTGLTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSTPTTMNYNQFLEKTKQLRLDQDVYGNATYFNENGRVYDTGNSWSEV 480
Db 421 LNAQKDFSTPTTMNYNQFLEKTKQLRLDQDVYGNATYFNENGRVYDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPKNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPKNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKQKQLAELNATNIYTVLDKIKLAKMILIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKQKQLAELNATNIYTVLDKIKLAKMILIRDKRFHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTGLLNLNDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLNLNDKDIRKILSGYVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKNDKPLVYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKNDKPLVYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 2

US-09-848-909-1

; Sequence 1, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-1

Query Match

Best Local Similarity 99.7%; Score 3761; DB 12; Length 736;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLLNESESSQGLLYFFDLNFAQPMVVTSTTGDLSPSELENIPSENOYF 60
Db 1 EVKQENRLLNESESSQGLLYFFDLNFAQPMVVTSTTGDLSPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDBYEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDBYEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNQLQPLKQKSSNSRKRSTSGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNQLQPLKQKSSNSRKRSTSGPTVPDRDN 180
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSSTASDYSDFEKT 240
Db 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSSTASDYSDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHAGFFDIGSVSAGFSNSSTVAIDHSLSLAGERTWAGTGLTADTARL 360
Db 301 SEVHGNAEVHAGFFDIGSVSAGFSNSSTVAIDHSLSLAGERTWAGTGLTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSTPTTMNYNQFLEKTKQLRLDQDVYGNATYFNENGRVYDTGNSWSEV 480
Db 421 LNAQKDFSTPTTMNYNQFLEKTKQLRLDQDVYGNATYFNENGRVYDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPKNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPKNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKQKQLAELNATNIYTVLDKIKLAKMILIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKQKQLAELNATNIYTVLDKIKLAKMILIRDKRFHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTGLLNLNDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLNLNDKDIRKILSGYVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKNDKPLVYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKNDKPLVYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 3

US-09-848-909-2

; Sequence 2, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-2

```
Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENLLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
DB 1 EVKQENLLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDAHNVMTWDDQEVINKASNKIRLEKGRLYQIKIQY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDAHNVMTWDDQEVINKASNKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNQLPELKQKSSNRKRSKSTASGPTVDPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNQLPELKQKSSNRKRSKSTASGPTVDPDRN 180
QY 181 DGI PDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTADPPYDFEKT 240
DB 181 DGI PDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTADPPYDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPARHPLVAAYPVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGSPNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGSPNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNPNYPKSLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNPNYPKSLAPIA 420
QY 421 LNAQKDFSSPTITWNYNQFLEKTKQLRLDQVYGNIAATYPENGRVVDTSNWSSEV 480
DB 421 LNAQKDFSSPTITWNYNQFLEKTKQLRLDQVYGNIAATYPENGRVVDTSNWSSEV 480
QY 481 LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGL 540
DB 481 LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGL 540
QY 541 QYQKXDIETFPDFNQDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKMILIRDKRPHYDR 600
DB 541 QYQKXDIETFPDFNQDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKMILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
DB 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNDKPLPLYSNPKNVYAVTKENTINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKKYNDKPLPLYSNPKNVYAVTKENTINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735
```

RESULT 4

```
US-09-848-909-3
; Sequence 3, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; OF BACTERIAL INFECTION AND PREVENTION OF BACTERIAL INFECTION
; FILE REFERENCE: 0074/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
```

```
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-09-848-909-3
```

Query Match 99.7%; Score 3761; DB 12; Length 736;

Best Local Similarity 99.7%; Pred. No. 4.3e-272;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 EVKQENLLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
DB 1 EVKQENLLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDAHNVMTWDDQEVINKASNKIRLEKGRLYQIKIQY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDAHNVMTWDDQEVINKASNKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNQLPELKQKSSNRKRSKSTASGPTVDPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNQLPELKQKSSNRKRSKSTASGPTVDPDRN 180
QY 181 DGI PDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTADPPYDFEKT 240
DB 181 DGI PDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTADPPYDFEKT 240
QY 241 GRIDKNVSPARHPLVAAYPVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPARHPLVAAYPVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGSPNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGSPNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNPNYPKSLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNPNYPKSLAPIA 420
QY 421 LNAQKDFSSPTITWNYNQFLEKTKQLRLDQVYGNIAATYPENGRVVDTSNWSSEV 480
DB 421 LNAQKDFSSPTITWNYNQFLEKTKQLRLDQVYGNIAATYPENGRVVDTSNWSSEV 480
QY 481 LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGL 540
DB 481 LPOIQTETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGL 540
QY 541 QYQKXDIETFPDFNQDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKMILIRDKRPHYDR 600
DB 541 QYQKXDIETFPDFNQDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKMILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
DB 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNDKPLPLYSNPKNVYAVTKENTINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKKYNDKPLPLYSNPKNVYAVTKENTINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735
```

RESULT 5

```
US-09-848-909-4
; Sequence 4, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
```

```
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-4
```

```
Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYFFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFTKVKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFTKVKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTSGAGTVPDRDN 180
Db 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTSGAGTVPDRDN 180
QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASPYSDFEKVT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASPYSDFEKVT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
QY 361 NANIRVNTGTAPIYNNVPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYNNVPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGFNEPENG 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGFNEPENG 540
QY 541 QYQKDIITBFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKNAKXNMLIRDKRPHYDR 600
Db 541 QYQKDIITBFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKNAKXNMLIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTGLLINDKIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLINDKIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
QY 721 IKKILFSSKKGVEIG 735
Db 721 IKKILFSSKKGVEIG 735
```

RESULT 6

US-09-848-909-5

```
; Sequence 5, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-5
```

```
Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYFFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFTKVKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFTKVKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTSGAGTVPDRDN 180
Db 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTSGAGTVPDRDN 180
QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASPYSDFEKVT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASPYSDFEKVT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
QY 361 NANIRVNTGTAPIYNNVPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYNNVPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGFNEPENG 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFGFNEPENG 540
QY 541 QYQKDIITBFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKNAKXNMLIRDKRPHYDR 600
Db 541 QYQKDIITBFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKNAKXNMLIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTGLLINDKIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLINDKIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
QY 721 IKKILFSSKKGVEIG 735
```

```
Db 721 IKKILIFSKGYEIG 735
|||||
RESULT 7
US-09-848-909-6
; Sequence 6, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-6

Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTSGAGTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTSGAGTVPDRDN 180
Qy 181 DGIPTDSEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYDFEKT 240
Db 181 DGIPTDSEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSISLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSISLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSPSTPTMNYNQFLEKTKQLRLDTQVYGNITATYFNFNGRVRVDGTGSNWSEV 480
Db 421 LNAQKDFSPSTPTMNYNQFLEKTKQLRLDTQVYGNITATYFNFNGRVRVDGTGSNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNVERIAAVNPSPLETTKPDNTLKEALKIAPFNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERIAAVNPSPLETTKPDNTLKEALKIAPFNEPNGNL 540
Qy 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 600
Qy 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660
Qy 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660
```

```
Qy 661 DMLNISLRQDGKTFIDFKKYNKPLYISNPNYKVVVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISLRQDGKTFIDFKKYNKPLYISNPNYKVVVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 8
US-09-848-909-7
; Sequence 7, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-7

Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTSGAGTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTSGAGTVPDRDN 180
Qy 181 DGIPTDSEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYDFEKT 240
Db 181 DGIPTDSEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSISLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSISLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSPSTPTMNYNQFLEKTKQLRLDTQVYGNITATYFNFNGRVRVDGTGSNWSEV 480
Db 421 LNAQKDFSPSTPTMNYNQFLEKTKQLRLDTQVYGNITATYFNFNGRVRVDGTGSNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNVERIAAVNPSPLETTKPDNTLKEALKIAPFNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERIAAVNPSPLETTKPDNTLKEALKIAPFNEPNGNL 540
Qy 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 600
```

Db 541 QYQKLTDFDFNDOQTQNIKNQALNATNIYTVLKDILNANMILIRDRPHYDR 600
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVNVAVTKENTIINPSENGDSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVNVAVTKENTIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
RESULT 9
US-09-848-909-8
; Sequence 8, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-8

Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENRLLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKGLDPLKYWTDSQNKKEVSSDNQLQPELKQSSNSRKRSTAGTPVDRDN 180
Db 121 QRENPTKGLDPLKYWTDSQNKKEVSSDNQLQPELKQSSNSRKRSTAGTPVDRDN 180
QY 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSTRITSKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSTRITSKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 420
QY 421 LNAQKQFSSTPTIMVNYQFIELEKTKQLRLDITDQVYGNITATYFNFGVRVDTGNSWSEV 480
Db 421 LNAQKQFSSTPTIMVNYQFIELEKTKQLRLDITDQVYGNITATYFNFGVRVDTGNSWSEV 480

QY 481 LPQIGETTARIIFNGKOLNVERRIAANVPSDPLETTKPDWTLKEALKIAPGNEPNGL 540
Db 481 LPQIGETTARIIFNGKOLNVERRIAANVPSDPLETTKPDWTLKEALKIAPGNEPNGL 540
QY 541 QYQKDIITFEFDFNFQOOTSQNIKNQALNATNIYTVLKDILNANMILIRDRPHYDR 600
Db 541 QYQKDIITFEFDFNFQOOTSQNIKNQALNATNIYTVLKDILNANMILIRDRPHYDR 600
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVNVAVTKENTIINPSENGDSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVNVAVTKENTIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 10
US-09-848-909-9
; Sequence 9, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-9

Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENRLLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120
QY 121 QRENPTKGLDPLKYWTDSQNKKEVSSDNQLQPELKQSSNSRKRSTAGTPVDRDN 180
Db 121 QRENPTKGLDPLKYWTDSQNKKEVSSDNQLQPELKQSSNSRKRSTAGTPVDRDN 180
QY 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSTRITSKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSTRITSKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 420

```
Db 361 NANIRYVNTGTAPIYNNVLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Qy 481 LPOIQTETARIIPNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIPNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
Qy 541 QYQKDIETEDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNKAKMILLIRKRFHYDR 600
Db 541 QYQKDIETEDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNKAKMILLIRKRFHYDR 600
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISSLRODGTDFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGTDFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 11
US-09-848-909-10
; Sequence 10, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 736
; TYPE: PR1
; ORGANISM: Bacillus anthracis
US-09-848-909-10

Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
Qy 241 GRIDKNVSPEARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSTRISKVTSRHT 300
Db 241 GRIDKNVSPEARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSTRISKVTSRHT 300
```

```
Qy 301 SEVHGNAEVHASFDFIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDFIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYNNVLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Qy 481 LPOIQTETARIIPNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIPNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
Qy 541 QYQKDIETEDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNKAKMILLIRKRFHYDR 600
Db 541 QYQKDIETEDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNKAKMILLIRKRFHYDR 600
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISSLRODGTDFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGTDFDKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 12
US-09-848-909-11
; Sequence 11, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 736
; TYPE: PR1
; ORGANISM: Bacillus anthracis
US-09-848-909-11

Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
```

Db 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKASTASDPYSDFEKT 240
Qy 241 GRIDKNVSPPEARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 SEVHGNAEVAHAFDILGGVSAGFSNSSTVAIDHSLSLAGERWAEWTMGNTADTARL 360
Db 301 SEVHGNAEVAHAFDILGGVSAGFSNSSTVAIDHSLSLAGERWAEWTMGNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540
Qy 541 QYQKDIITFDNFDOOTSONIKQLAELNATNIYVLDKIKLNAMNILLIRDKRPHYDR 600
Db 541 QYQKDIITFDNFDOOTSONIKQLAELNATNIYVLDKIKLNAMNILLIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAAREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAAREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFDFKXNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFDFKXNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 13

US-09-848-909-12

; Sequence 12, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; PRIOR FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-12

Query Match

Best Local Similarity 99.7%; Score 3761; DB 12; Length 736;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFFDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESQGLGYFFDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYFATSDADNHVTWVDQEVINKANSNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYFATSDADNHVTWVDQEVINKANSNKIRLEKGLYQIKIY 120

Qy 121 QRENPTKEGLDFKLYNTDSQNKKEVISSDNLQLPKQKSSNRKCRSTSAGTVPDRDN 180
Db 121 QRENPTKEGLDFKLYNTDSQNKKEVISSDNLQLPKQKSSNRKCRSTSAGTVPDRDN 180
Qy 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKASTASDPYSDFEKT 240
Db 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKASTASDPYSDFEKT 240
Qy 241 GRIDKNVSPPEARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Qy 301 SEVHGNAEVAHAFDILGGVSAGFSNSSTVAIDHSLSLAGERWAEWTMGNTADTARL 360
Db 301 SEVHGNAEVAHAFDILGGVSAGFSNSSTVAIDHSLSLAGERWAEWTMGNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540
Qy 541 QYQKDIITFDNFDOOTSONIKQLAELNATNIYVLDKIKLNAMNILLIRDKRPHYDR 600
Db 541 QYQKDIITFDNFDOOTSONIKQLAELNATNIYVLDKIKLNAMNILLIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAAREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAAREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFDFKXNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFDFKXNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 14

US-09-848-909-13

; Sequence 13, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; PRIOR FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-13

Query Match

Best Local Similarity 99.7%; Score 3761; DB 12; Length 736;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFFDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60


```
Db 1 EVKQENRLLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKEGLDPLKLYWTDSONKKEVTSNDNLQPLKOKSSNSRKRSTASAGTVPDRDN 180
Db 121 QRENPTKEGLDPLKLYWTDSONKKEVTSNDNLQPLKOKSSNSRKRSTASAGTVPDRDN 180
QY 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYSDFE 240
Db 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYSDFE 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERETWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERETWAGTGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYVNFQLELEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITWYVNFQLELEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480
QY 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFDOOTSONIKNQLAELNATNIYTVLDKIKLNKAMNIIIRDKEFHYDR 600
Db 541 QYQKDIITEFDNFDOOTSONIKNQLAELNATNIYTVLDKIKLNKAMNIIIRDKEFHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRODGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

RESULT 15

```
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/06002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
```

US-09-848-909-14

```
Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4,3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENRLLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKEGLDPLKLYWTDSONKKEVTSNDNLQPLKOKSSNSRKRSTASAGTVPDRDN 180
Db 121 QRENPTKEGLDPLKLYWTDSONKKEVTSNDNLQPLKOKSSNSRKRSTASAGTVPDRDN 180
QY 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYSDFE 240
Db 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYSDFE 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERETWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERETWAGTGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYVNFQLELEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITWYVNFQLELEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480
QY 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFDOOTSONIKNQLAELNATNIYTVLDKIKLNKAMNIIIRDKEFHYDR 600
Db 541 QYQKDIITEFDNFDOOTSONIKNQLAELNATNIYTVLDKIKLNKAMNIIIRDKEFHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRODGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

Search completed: May 3, 2004, 20:02:09
Job time : 41.6393 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 175.546 Seconds
(without alignments)
4086.665 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKLIFSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/2/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/2/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/2/paa/US107_COMB.pep.*
33: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

	3774	100.0	735	1	PCT-US03-35733-10	Sequence 10, Appl
1	3774	100.0	735	23	US-09-848-909A-10	Sequence 10, Appl
2	3771	99.9	735	1	PCT-US03-19786-4	Sequence 4, Appl
3	3768	99.8	735	1	PCT-US03-35733-2	Sequence 2, Appl
4	3768	99.8	735	23	US-09-848-909A-2	Sequence 2, Appl
5	3767	99.8	735	1	PCT-US03-35733-8	Sequence 8, Appl
6	3767	99.8	735	23	US-09-848-909A-8	Sequence 8, Appl
7	3766	99.8	735	1	PCT-US03-35733-13	Sequence 13, Appl
8	3766	99.8	735	23	US-09-848-909A-13	Sequence 13, Appl
9	3763	99.7	735	1	PCT-US03-35733-7	Sequence 7, Appl
10	3763	99.7	735	23	US-09-848-909A-7	Sequence 7, Appl
11	3762	99.7	735	1	PCT-US03-35733-4	Sequence 4, Appl
12	3762	99.7	735	23	US-09-848-909A-4	Sequence 4, Appl
13	3762	99.7	735	23	US-09-848-909A-6	Sequence 6, Appl
14	3762	99.7	735	1	PCT-US03-35733-5	Sequence 5, Appl
15	3761	99.7	735	1	PCT-US03-35733-11	Sequence 11, Appl
16	3761	99.7	735	23	US-09-848-909A-11	Sequence 11, Appl
17	3761	99.7	735	1	PCT-US03-35733-19	Sequence 19, Appl
18	3761	99.7	735	23	US-09-848-909A-19	Sequence 19, Appl
19	3761	99.7	735	1	PCT-US03-35733-20	Sequence 20, Appl
20	3761	99.7	735	1	PCT-US03-35733-21	Sequence 21, Appl
21	3761	99.7	735	22	US-09-791-537-43735	Sequence 43735, A
22	3761	99.7	735	23	US-09-848-909A-5	Sequence 5, Appl
23	3761	99.7	735	23	US-09-848-909A-11	Sequence 11, Appl
24	3761	99.7	735	23	US-09-848-909A-19	Sequence 19, Appl
25	3761	99.7	735	23	US-09-848-909A-20	Sequence 20, Appl
26	3761	99.7	735	23	US-09-848-909A-21	Sequence 21, Appl
27	3761	99.7	735	30	US-10-410-647-30	Sequence 30, Appl
28	3761	99.7	736	1	PCT-US01-14372A-1	Sequence 1, Appl
29	3761	99.7	736	1	PCT-US01-14372A-2	Sequence 2, Appl
30	3761	99.7	736	1	PCT-US01-14372A-3	Sequence 3, Appl
31	3761	99.7	736	1	PCT-US01-14372A-4	Sequence 4, Appl
32	3761	99.7	736	1	PCT-US01-14372A-5	Sequence 5, Appl
33	3761	99.7	736	1	PCT-US01-14372A-6	Sequence 6, Appl
34	3761	99.7	736	1	PCT-US01-14372A-7	Sequence 7, Appl
35	3761	99.7	736	1	PCT-US01-14372A-8	Sequence 8, Appl
36	3761	99.7	736	1	PCT-US01-14372A-9	Sequence 9, Appl
37	3761	99.7	736	1	PCT-US01-14372A-10	Sequence 10, Appl
38	3761	99.7	736	1	PCT-US01-14372A-11	Sequence 11, Appl
39	3761	99.7	736	1	PCT-US01-14372A-12	Sequence 12, Appl
40	3761	99.7	736	1	PCT-US01-14372A-13	Sequence 13, Appl
41	3761	99.7	736	1	PCT-US01-14372A-14	Sequence 14, Appl
42	3761	99.7	736	1	PCT-US01-14372A-15	Sequence 15, Appl
43	3761	99.7	736	1	PCT-US01-14372A-16	Sequence 16, Appl
44	3761	99.7	736	1	PCT-US01-14372A-17	Sequence 17, Appl
45	3761	99.7	736	1	PCT-US01-14372A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
PCT-US03-35733-10
; Sequence 10, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-10

Query Match 100.0%; Score 3774; DB 1; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.le-299;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGFFDLNFAQPMVVTSTTGDLSIPSSSELENIPISENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFAQPMVVTSTTGDLSIPSSSELENIPISENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNQLPELKOKSSNSRKRSTASGPTVPDRN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNQLPELKOKSSNSRKRSTASGPTVPDRN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEYV 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEYV 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
QY 301 SEVHGNAEVAHAFDIDGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Db 301 SEVHGNAEVAHAFDIDGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSTPTIMYNNQFLEKTKQLRLDQVYGNIAIYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDFSTPTIMYNNQFLEKTKQLRLDQVYGNIAIYNFENGVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMLKEALKIAFGFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMLKEALKIAFGFNEPNGNL 540
QY 541 QYQKDIITFDFNFDOQTSONIKQLAELNATNIYTVLDKIKLNAXMILIRDKRPHYDR 600
Db 541 QYQKDIITFDFNFDOQTSONIKQLAELNATNIYTVLDKIKLNAXMILIRDKRPHYDR 600
QY 601 NNIAVGADESVMKEAHREVINGSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHREVINGSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRODQKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDSTNG 720
Db 661 DMLNSSLRODQKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 2

US-09-848-909A-10
; Sequence 10, Application US/09848909A

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909A

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-05-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

; TYPE: PRT

US-09-848-909A-10

Query Match 100.0%; Score 3774; DB 23; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.1e-299;

Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGFFDLNFAQPMVVTSTTGDLSIPSSSELENIPISENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFAQPMVVTSTTGDLSIPSSSELENIPISENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNQLPELKOKSSNSRKRSTASGPTVPDRN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNQLPELKOKSSNSRKRSTASGPTVPDRN 180
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEYV 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEYV 240
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300
QY 301 SEVHGNAEVAHAFDIDGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Db 301 SEVHGNAEVAHAFDIDGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSTPTIMYNNQFLEKTKQLRLDQVYGNIAIYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDFSTPTIMYNNQFLEKTKQLRLDQVYGNIAIYNFENGVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMLKEALKIAFGFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMLKEALKIAFGFNEPNGNL 540
QY 541 QYQKDIITFDFNFDOQTSONIKQLAELNATNIYTVLDKIKLNAXMILIRDKRPHYDR 600
Db 541 QYQKDIITFDFNFDOQTSONIKQLAELNATNIYTVLDKIKLNAXMILIRDKRPHYDR 600
QY 601 NNIAVGADESVMKEAHREVINGSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHREVINGSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRODQKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDSTNG 720
Db 661 DMLNSSLRODQKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 3

PCT-US03-19786-4

; Sequence 4, Application PC/TUS0319786

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.

; TITLE OF INVENTION: Modified Shine Dalgarno Sequences and Methods of Use Thereof

; FILE REFERENCE: PV595PCT

; CURRENT APPLICATION NUMBER: PCT/US03/19786

; CURRENT FILING DATE: 2003-06-25

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 735

; TYPE: PRT

; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-2

Query Match 99.8%; Score 3768; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 3.3e-299;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60
QY 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKOKSSNSRKRSTAGTVPDRDN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKOKSSNSRKRSTAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
DB 181 DGIPDSLEVEGYTVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTHT 300
QY 301 SEVHGNAEYHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEYHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRDLTDQVYGNIAATYNFENGVRVDTGSNWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRDLTDQVYGNIAATYNFENGVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKONILLIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKONILLIRDKRPHYDR 600
QY 601 NNTAVGADESVMKEAHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNTAVGADESVMKEAHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRODGKTFIDFKKYNKPLIYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
DB 661 DMLNSSLRODGKTFIDFKKYNKPLIYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 6
PCT-US03-35733-8

; Sequence 8, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-8

Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 3.9e-299;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60
QY 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKOKSSNSRKRSTAGTVPDRDN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKOKSSNSRKRSTAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
DB 181 DGIPDSLEVEGYTVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTHT 300
QY 301 SEVHGNAEYHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEYHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRDLTDQVYGNIAATYNFENGVRVDTGSNWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRDLTDQVYGNIAATYNFENGVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKONILLIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKONILLIRDKRPHYDR 600
QY 601 NNTAVGADESVMKEAHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNTAVGADESVMKEAHREVINSTEGLLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRODGKTFIDFKKYNKPLIYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
DB 661 DMLNSSLRODGKTFIDFKKYNKPLIYISNPNYKVNVAATKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

Db 661 DMLNLSLRQDGKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720

QY 721 IKKILIFSCKGYEIG 735

Db 721 IKKILIFSCKGYEIG 735

RESULT 8

PCT-US03-35733-13

; Sequence 13, Application PC/TUS0335733

; GENERAL INFORMATION:

; APPLICANT: President and Fellows of Harvard College et al.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/072003

; CURRENT APPLICATION NUMBER: PCT/US03/35733

; CURRENT FILING DATE: 2003-10-10

; PRIOR APPLICATION NUMBER: US 60/424,987

; PRIOR FILING DATE: 2002-11-08

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

PCT-US03-35733-13

Query Match 99.8%; Score 3766; DB 1; Length 735;

Best Local Similarity 99.9%; Pred. No. 4,8e-299;

Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTSGAGTVPDRDN 180

Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTSGAGTVPDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYSDFEKT 240

Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYSDFEKT 240

QY 241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSSTRTHT 300

Db 241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSSTRTHT 300

QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYNNVLTSLVLGKQTLATIKADENQSLQILAPNNYPSKQLAPIA 420

Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKQTLATIKADENQSLQILAPNNYPSKQLAPIA 420

QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTQVYGNATYFNENGRVRVDTGSNWSEV 480

Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTQVYGNATYFNENGRVRVDTGSNWSEV 480

QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

QY 541 QYQKDIETEDFDFDQSTQSNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600

Db 541 QYQKDIETEDFDFDQSTQSNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600

QY 601 NNIAVGADES VVKAHREVNINSSTEGLLNIDKIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

US-09-848-909A-8

; Sequence 8, Application US/09848909A

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909A

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-05-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909A-8

Query Match 99.8%; Score 3767; DB 23; Length 735;

Best Local Similarity 99.9%; Pred. No. 3.9e-299;

Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTSGAGTVPDRDN 180

Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTSGAGTVPDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYSDFEKT 240

Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYSDFEKT 240

QY 241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSSTRTHT 300

Db 241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSSTRTHT 300

QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYNNVLTSLVLGKQTLATIKADENQSLQILAPNNYPSKQLAPIA 420

Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKQTLATIKADENQSLQILAPNNYPSKQLAPIA 420

QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTQVYGNATYFNENGRVRVDTGSNWSEV 480

Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTQVYGNATYFNENGRVRVDTGSNWSEV 480

QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

QY 541 QYQKDIETEDFDFDQSTQSNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600

Db 541 QYQKDIETEDFDFDQSTQSNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600

QY 601 NNIAVGADES VVKAHREVNINSSTEGLLNIDKIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

Db 601 NNIAVGADES VVKAHREVNINSSTEGLLNIDKIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

QY 661 DMLNLSLRQDGKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720

Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVNVAVTKEINTIINPSENGDTSNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVNVAVTKEINTIINPSENGDTSNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 9
US-09-848-909A-13
; Sequence 13, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; PRIOR FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-13

Query Match 99.8%; Score 3766; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 4.8e-299;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSSITGDLSPSSSELENIPESENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSSITGDLSPSSSELENIPESENQYF 60

Qy 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKSRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKSRSTASGTPVDRDN 180

Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

Qy 421 LNAQKDFSSPTIMNYNQFLEKTKQLRLDTPQVYGNATATYNFENGRVVDTSNWSSEV 480
Db 421 LNAQKDFSSPTIMNYNQFLEKTKQLRLDTPQVYGNATATYNFENGRVVDTSNWSSEV 480

Qy 481 LPOIQTETARIIFNGKOLNVERIAAVNPSPDELTTPKDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIFNGKOLNVERIAAVNPSPDELTTPKDMTLKEALKIAGFNEPENG 540

Qy 541 QYQKDI TEFPNFDQOTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRFFHYDR 600

Db 541 QYQKDI TEFPNFDQOTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRFFHYDR 600
Qy 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVNVAVTKEINTIINPSENGDTSNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVNVAVTKEINTIINPSENGDTSNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 10
PCT-US03-35733-7
; Sequence 7, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; PRIOR FILING DATE: 2003-10-10
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-7

Query Match 99.7%; Score 3763; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 8.4e-299;
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSSITGDLSPSSSELENIPESENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSSITGDLSPSSSELENIPESENQYF 60

Qy 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKSRSTASGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKSRSTASGTPVDRDN 180

Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

Qy 421 LNAQKDFSSPTIMNYNQFLEKTKQLRLDTPQVYGNATATYNFENGRVVDTSNWSSEV 480
Db 421 LNAQKDFSSPTIMNYNQFLEKTKQLRLDTPQVYGNATATYNFENGRVVDTSNWSSEV 480

Qy 481 LPOIQTETARIIFNGKOLNVERIAAVNPSPDELTTPKDMTLKEALKIAGFNEPENG 540

Db 481 LPOIETTARIIFNGKOLNVERRIAANVPSDPLETTKPDMLKALKIAFGNEPENG 540
Qy 541 QYCGKDIITFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKWNILIRDRKFHYDR 600
Db 541 QYCGKDIITFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKWNILIRDRKFHYDR 600
Qy 601 NNTAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNTAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRODQGTFTDFKKNNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRODQGTFTDFKKNNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 11

US-09-848-909A-7

; Sequence 7, Application US/09848909A

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909A

; PRIOR FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-05-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909A-7

Query Match 99.7%; Score 3763; DB 23; Length 735;
Best Local Similarity 99.7%; Pred. No. 8.4e-299;
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNRKSRKSTAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNRKSRKSTAGTPVDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWSSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWSSTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Qy 361 NANRYNVTGTAPIYVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANRYNVTGTAPIYVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420

Qy 421 LNAQKDPSTPTIMYNOFLEKTKQRLDQVYGNIAFYENFENGVRVDTGNSMSEV 480
Db 421 LNAQKDPSTPTIMYNOFLEKTKQRLDQVYGNIAFYENFENGVRVDTGNSMSEV 480
Qy 481 LPOIETTARIIFNGKOLNVERRIAANVPSDPLETTKPDMLKALKIAFGNEPENG 540
Db 481 LPOIETTARIIFNGKOLNVERRIAANVPSDPLETTKPDMLKALKIAFGNEPENG 540
Qy 541 QYCGKDIITFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKWNILIRDRKFHYDR 600
Db 541 QYCGKDIITFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKWNILIRDRKFHYDR 600
Qy 601 NNTAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNTAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRODQGTFTDFKKNNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRODQGTFTDFKKNNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 12

PCT-US03-35733-4

; Sequence 4, Application PC/TUS03035733

; GENERAL INFORMATION:

; APPLICANT: President and Fellows of Harvard College et al.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/072003

; CURRENT APPLICATION NUMBER: PCT/US03/35733

; CURRENT FILING DATE: 2003-10-10

; PRIOR APPLICATION NUMBER: US 60/424,987

; PRIOR FILING DATE: 2002-11-08

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

PCT-US03-35733-4

Query Match 99.7%; Score 3762; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 1e-298;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNRKSRKSTAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNRKSRKSTAGTPVDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWSSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWSSTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360

```

QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKNOTLATIKADENQLSQILAPNNYPSKULAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVGLKNOTLATIKADENQLSQILAPNNYPSKULAPIA 420
QY 421 LNAQKDFSSPTITMNYNOFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
DB 421 LNAQKDFSSPTITMNYNOFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENG 540
QY 541 QYQKDIETEDFNFDOOTSNIKNQLAELNATNIYVLDKIKLNAKMLIRDKRPFHYDR 600
DB 541 QYQKDIETEDFNFDOOTSNIKNQLAELNATNIYVLDKIKLNAKMLIRDKRPFHYDR 600
QY 601 NNTAVGADESUVKEAHREVINSTEGLLNIDKDIRKILSGYIVIEBTEGLKEVINDRY 660
DB 601 NNTAVGADESUVKEAHREVINSTEGLLNIDKDIRKILSGYIVIEBTEGLKEVINDRY 660
QY 661 DMLNISSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNISSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
DB 721 IKKILIFSKGYEIG 735

```

RESULT 13

PCT-US03-35733-6

; Sequence 6, Application PC/TUS0335733

; GENERAL INFORMATION:

; APPLICANT: President and Fellows of Harvard College et al.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE REFERENCE: 00742/072003

; CURRENT APPLICATION NUMBER: PCT/US03/35733

; PRIOR FILING DATE: 2003-10-10

; PRIOR APPLICATION NUMBER: US 60/424,987

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; TYPE: PRT

; ORGANISM: Bacillus anthracis

PCT-US03-35733-6

Query Match 99.7%; Score 3762; DB 1; Length 735;

Best Local Similarity 99.7%; Pred. No. 1e-298;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSGDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSGDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKRSSTAGTVPDRDN 180
QY 181 DGIIPDSLEVEGYTVDVNKRFTFSPWISNIHEKKGLTKYKSPKWTASDPSPEKVT 240
DB 181 DGIIPDSLEVEGYTVDVNKRFTFSPWISNIHEKKGLTKYKSPKWTASDPSPEKVT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSEPTIKNTSRTHT 300
DB 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSEPTIKNTSRTHT 300

```

```

QY 301 SEVHGNAEVHASFDFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKNOTLATIKADENQLSQILAPNNYPSKULAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVGLKNOTLATIKADENQLSQILAPNNYPSKULAPIA 420
QY 421 LNAQKDFSSPTITMNYNOFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
DB 421 LNAQKDFSSPTITMNYNOFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPENG 540
QY 541 QYQKDIETEDFNFDOOTSNIKNQLAELNATNIYVLDKIKLNAKMLIRDKRPFHYDR 600
DB 541 QYQKDIETEDFNFDOOTSNIKNQLAELNATNIYVLDKIKLNAKMLIRDKRPFHYDR 600
QY 601 NNTAVGADESUVKEAHREVINSTEGLLNIDKDIRKILSGYIVIEBTEGLKEVINDRY 660
DB 601 NNTAVGADESUVKEAHREVINSTEGLLNIDKDIRKILSGYIVIEBTEGLKEVINDRY 660
QY 661 DMLNISSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNISSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKGYEIG 735
DB 721 IKKILIFSKGYEIG 735

```

RESULT 14

US-09-848-909A-4

; Sequence 4, Application US/09848909A

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909A

; PRIOR FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909A-4

Query Match 99.7%; Score 3762; DB 23; Length 735;

Best Local Similarity 99.7%; Pred. No. 1e-298;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSGDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSGDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKRSSTAGTVPDRDN 180
QY 181 DGIIPDSLEVEGYTVDVNKRFTFSPWISNIHEKKGLTKYKSPKWTASDPSPEKVT 240
DB 181 DGIIPDSLEVEGYTVDVNKRFTFSPWISNIHEKKGLTKYKSPKWTASDPSPEKVT 240

```

QY	241	GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTSRHT	300
Db	241	GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTSRHT	300
QY	301	SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTGLNTADTARL	360
Db	301	SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTGLNTADTARL	360
QY	361	NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
Db	361	NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVNFENGVRVDTGNSWSEV	480
Db	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVNFENGVRVDTGNSWSEV	480
QY	481	LPQIQETTARIIFNGKDLNVERIAAVNPSPDLETTKPDMLKEALKIAFGNEPNGNL	540
Db	481	LPQIQETTARIIFNGKDLNVERIAAVNPSPDLETTKPDMLKEALKIAFGNEPNGNL	540
QY	541	QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNMLIRDKRPHYDR	600
Db	541	QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNMLIRDKRPHYDR	600
QY	601	NNIAVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY	660
Db	601	NNIAVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY	660
QY	661	DMLNISSLRQDGKTFIDFKYNDKLPYISPNYKVNVAVTKENTIINPSENGDTSTNG	720
Db	661	DMLNISSLRQDGKTFIDFKYNDKLPYISPNYKVNVAVTKENTIINPSENGDTSTNG	720
QY	721	IKKILIFSKKGYEIG 735	
Db	721	IKKILIFSKKGYEIG 735	
RESULT 15			
US-09-848-909A-6			
; Sequence 6, Application US/09848909A			
; GENERAL INFORMATION:			
; APPLICANT: Sellman, R. John			
; APPLICANT: Sellman, Brett R.			
; TITLE OF INVENTION: Compounds and Methods for the Treatment			
; TITLE OF INVENTION: and Prevention of Bacterial Infection			
; FILE REFERENCE: 00742/060002			
; CURRENT APPLICATION NUMBER: US/09/848,909A			
; CURRENT FILING DATE: 2001-05-04			
; PRIOR APPLICATION NUMBER: US 60/201,800			
; PRIOR FILING DATE: 2000-05-04			
; NUMBER OF SEQ ID NOS: 35			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 6			
; LENGTH: 735			
; TYPE: PRN			
; ORGANISM: Bacillus anthracis			
US-09-848-909A-6			
Query Match 99.7%; Score 3762; DB 23; Length 735;			
Best Local Similarity 99.7%; Pred. No. 1e-298;			
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	EVKQENRLNSESSESSQGLGYFFDLNFQAPMWVTSSTTGDLSPSSSELENIPSENQYF	60
Db	1	EVKQENRLNSESSESSQGLGYFFDLNFQAPMWVTSSTTGDLSPSSSELENIPSENQYF	60
QY	61	QSAIWSGFIKVKSDSYTTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY	120
Db	61	QSAIWSGFIKVKSDSYTTFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY	120
QY	121	ORENPTEKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTAGTVPDRDN	180

Db	121	ORENPTEKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTAGTVPDRDN	180
QY	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNIIHEKKGKLTXYKSSPEKMWSTASDPYSDFEKT	240
Db	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNIIHEKKGKLTXYKSSPEKMWSTASDPYSDFEKT	240
QY	241	GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTSRHT	300
Db	241	GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTSRHT	300
QY	301	SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTGLNTADTARL	360
Db	301	SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTGLNTADTARL	360
QY	361	NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
Db	361	NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVNFENGVRVDTGNSWSEV	480
Db	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVNFENGVRVDTGNSWSEV	480
QY	481	LPQIQETTARIIFNGKDLNVERIAAVNPSPDLETTKPDMLKEALKIAFGNEPNGNL	540
Db	481	LPQIQETTARIIFNGKDLNVERIAAVNPSPDLETTKPDMLKEALKIAFGNEPNGNL	540
QY	541	QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNMLIRDKRPHYDR	600
Db	541	QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNMLIRDKRPHYDR	600
QY	601	NNIAVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY	660
Db	601	NNIAVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY	660
QY	661	DMLNISSLRQDGKTFIDFKYNDKLPYISPNYKVNVAVTKENTIINPSENGDTSTNG	720
Db	661	DMLNISSLRQDGKTFIDFKYNDKLPYISPNYKVNVAVTKENTIINPSENGDTSTNG	720
QY	721	IKKILIFSKKGYEIG 735	
Db	721	IKKILIFSKKGYEIG 735	

Search completed: May 3, 2004, 19:57:40
Job time : 178.546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 13.0855 Seconds
(without alignments)

5403.004 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENLLNESSSSQGL.....TSTNGIKLIFSKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3761	99.7	764	2 I39934	protective antigen
2	876	23.2	875	2 I40862	iota toxin compone
3	235.5	6.2	192	2 I39933	cryptic protein -
4	235.5	6.2	204	2 G59104	hypothetical prote
5	206	5.5	4688	2 F82885	hypothetical prote
6	188.5	5.0	2401	2 T28676	rhostry protein -
7	185.5	4.9	1639	2 S05603	major merozoite su
8	184.5	4.9	1125	2 E95998	membrane nucleas
9	184.5	4.9	2529	2 B84635	toxin-like outer m
10	182	4.8	6713	2 B89921	hypothetical prote
11	181	4.8	2869	2 T28677	rhostry protein -
12	181	4.8	4152	2 T31102	filamentous hemagg
13	180	4.8	1193	2 S68218	botulinum neurotox
14	179.5	4.8	1072	2 A86827	hypothetical prote
15	179	4.7	1635	2 A10452	hemolysin [importe
16	178	4.7	1227	2 C97033	uncharacterized pr
17	177	4.7	752	2 G30599	hypothetical prote
18	177	4.7	1302	1 JC6009	surface-located me
19	177	4.7	4919	2 T31105	hypothetical prote
20	174.5	4.6	1365	2 T30822	lmp1 protein - Myc
21	174	4.6	1658	2 S55101	repeat organellar
22	173	4.6	1939	2 T18372	hypothetical prote
23	173	4.6	5005	2 F82884	hypothetical prote
24	172.5	4.5	2399	2 H71879	toxin-like outer m
25	171.5	4.5	821	2 S67087	hypothetical prote
26	171	4.5	2178	2 S55805	alpha-toxin - C1os
27	170	4.5	1631	1 SAZQK1	major merozoite su
28	170	4.5	1837	2 T41023	probable nuclear p
29	169.5	4.5	1230	2 S56850	SMC1 protein homol

ALIGNMENTS

RESULT 1

I39934
protective antigen precursor - Bacillus anthracis plasmid
C/Species: Bacillus anthracis
C/Date: 19-Jul-1996 #sequence revision 19-Jul-1996 #text_change 01-Dec-2000
C/Accession: I39934; S89160; F59104
R/Weikos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A/Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr.
A/Reference number: I39933; MUID:89172073; PMID:3148491
A/Accession: I39934
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-764 <RES>
A/Cross-references: GB:M2589; NID:9143280; PIDN:AAA22637.1; PID:G143282
R/Friedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klmpel, K.R.; Birch, N.P.; Loh, Y.P.
Arch. Biochem. Biophys. 316, 5-13, 1995
A/Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SP
A/Reference number: S69160; MUID:95142670; PMID:7840657
A/Accession: S69160
A/Molecule type: protein
A/Residues: 197-202 <FRI>
R/Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle
J. Bacteriol. 181, 6509-6515, 1999
A/Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbor
A/Reference number: A59091; MUID:99445483; PMID:10515943
A/Accession: F59104
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-313, 'Q', 315-764 <OKI>
A/Cross-references: GB:AF065404; NID:G4894216; PIDN:AAD32414.1; PID:G4894326
A/Experimental source: strain Sterne
A/Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid
C/Genetics:
A/Gene: pXO1-110
A/Genome: plasmid
C/Function:
C/Description: three component exotoxin; protective antigen binds to receptors on the s
Y active components edema factor or lethal factor; the complex is internalized by recep
C/Keywords: exotoxin
F/1-29/Domain: signal sequence #status predicted <SIG>
F/30-156/Domain: propeptide #status predicted <PRO>
F/197-202/Product: protective antigen #status experimental <MAT>

Query Match 99.7%; Score 3761; DB 2; Length 764;

Best Local Similarity 99.7%; Pred. No. 2.5e-182;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSSSQGLGYFSDLNFOAPMVVTSITGDLISPSSELENIPSENQYF 60

DB 30 EVKQENLLNESSSSQGLGYFSDLNFOAPMVVTSITGDLISPSSELENIPSENQYF 89

QY 587 MNLIIKDKPHFDRNNIAGAESVVKAEHREVINSSTEGLLNIDKIRKILSGYIVEI 646
 Db 1 MNLIVRDP-YHYDNNNGNIYGVDDSVLKAYKQILNWSGSLNLDVQALSGYMLQI 59
 QY 647 EDTE-----GLKEVINDRYDMLNIISSLRQDGKTFIDPKYNDKPLIYISNEN 693
 Db 60 KKPSEHLTNSPVTITLAGKDSGVGELYRVL-----DQAGFLDNKFDENWRSIV-DPG 112
 QY 694 YKVNVAVTKEN-TIINPSENGDTSTNGIKKILIFSKGYEI 734
 Db 113 DDVYVAVTKEDFNAVTRDENGNI-NKLKNTLVLSGKIKEI 153

RESULT 4
 G59104
 hypothetical protein pXOI-111 - Bacillus anthracis virulence plasmid pXOI
 C/Species: Bacillus anthracis
 C/Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-May-2000
 C/Accession: G59104
 R/Okinaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
 J. Bacteriol. 181, 6509-6515, 1999
 A/Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbored
 A/Reference number: A59091, MUID:99445483; PMID:10515943
 A/Accession: G59104
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-204 <OKI>
 A/Cross-references: GB:AF065404; NID:g4894216; PIDN:AA032415.1; PID:g4894327
 A/Experimental source: strain Sterne
 A/Note: similar to hypothetical protein in the protective antigen domain; ypa, plasmid p
 C/Genetics:
 A/Gene: pXOI-111
 A/Genome: plasmid

Query Match 6.2%; Score 235.5; DB 2; Length 204;
 Best Local Similarity 34.6%; Pred. No. 2.8e-05;
 Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNLIIKDKPHFDRNNIAGAESVVKAEHREVINSSTEGLLNIDKIRKILSGYIVEI 646
 Db 1 MNLIVRDP-YHYDNNNGNIYGVDDSVLKAYKQILNWSGSLNLDVQALSGYMLQI 59
 QY 647 EDTE-----GLKEVINDRYDMLNIISSLRQDGKTFIDPKYNDKPLIYISNEN 693
 Db 60 KKPSEHLTNSPVTITLAGKDSGVGELYRVL-----DQAGFLDNKFDENWRSIV-DPG 112
 QY 694 YKVNVAVTKEN-TIINPSENGDTSTNGIKKILIFSKGYEI 734
 Db 113 DDVYVAVTKEDFNAVTRDENGNI-NKLKNTLVLSGKIKEI 153

RESULT 5
 F82885
 hypothetical protein U0482 [imported] - Ureaplasma urealyticum
 C/Species: Ureaplasma urealyticum
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C/Accession: F82885
 R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to GenBank, February 2000
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
 A/Reference number: A82870
 A/Accession: F82885
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-468 <GL>
 A/Cross-references: GB:AE002145; GB:AF222894; NID:g56899476; PIDN:AAF30894.1; GSPDB:GN001
 A/Experimental source: serovar 3; biovar 1
 C/Genetics:
 A/Gene: U0482
 A/Genetic code: SGC3

Query Match 5.5%; Score 206; DB 2; Length 468;
 Best Local Similarity 21.3%; Pred. No. 0.09;

Matches 174; Conservative 132; Mismatches 324; Indels 186; Gaps 40;
 QY 19 LLGYFSLNFQAPMVTSSTTGDISIPSSSELENIPEENQVFOASIWGFIKVKKSDEYT 78
 Db 3699 LVDVYLD-NHQNIIDETKIFKQHN-V-SKEIENPGVTWISKGNWKSSTDTTANFEFK 3756
 QY 79 PATSADNHVWVDOEVINAKSNKIRLEKGLYQIKIQORENTEKGLDFKLTWD 138
 Db 3757 IETQ-----DDNDVLNIDATVRFKDEHNNIKQKIVRIKEN-----ND 3795
 QY 139 SQNKKEVSSDNLQPELKQSSN-----SRKGRSTAGTPVDRD----- 179
 Db 3796 WLKQOI---DNLN-PETKYKLENIELSKPLKTHNLVSINDKENSLITGPNVLKV 3851
 QY 180 -----NDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPSD 235
 Db 3852 IQQNDTINDTQQTINVTLSGVNSK-YNGROIKVYKNNNNVIYESS-----LITLQKQND 3907
 QY 236 FEKVTRGRIDKQVSPPEARPLVAAPIVHVDMENIILSKN-EDOSTQDTSRT-TISKNT 293
 Db 3908 YQLLSNLNSN-----REYRFEKIEINHSNTNNFEDLEKLVGNTFITOTKNT 3957
 QY 294 STSRTHS-EVHGNAEVAASF-----FDIGSVSAGPS-----NSNS---STVAID 335
 Db 3958 TVQNDSSATVGTGVNFNFKIKSEDKILENNOOVAVWAFAPKETIRDTTWLQYTRPLK 4017
 QY 336 HSLSLAGERTWAETMGLNT---ADTARLNANIRYVNTGTAPIVNVLPPTTSLVLGKQTLA 392
 Db 4018 DVTSPFKEGTWAHDLNSVNFKEETTYKLVKIQFVKNPTRKAKNNINNNVILDTNSI 4077
 QY 393 -----TIKADENQLSILAPNNYPPSKNLAPIALN-AQKDFSTPTWNY--NQFLELE 443
 Db 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQTINFTLSGVKKSWSGKKTKLSYKNDSESI 4137
 QY 444 KTKQLRLDTDOVYGNIAYNFENGK-----VRVDTGSNMSEVLPTQTOETARIFNGKD 497
 Db 4138 HNEVLIESNKTQYNILLNLLNLRKRTVTLIDVKLIDNNVSDPPKEGNTLNSFITRTSA 4197
 QY 498 LNLVEERAAVNPSPLETT-----KPDMTLKEA-----LKIAGFPEPENGLO 541
 Db 4198 INVLEIEISNRASNLKSTIIKINLNDPQVLEKQDQATIVYGNKKQAMGFTVSNIK 4257
 QY 542 YQKGDITDFDFDQDQTSQNTK-NQLAELNATNTYTVLDKIKLNKAKNIIIRDKRHYD- 599
 Db 4258 YLTATLVDLNFN-DKVNIVNISFNKKPSIAEN-----IGDKSNII-----YNNDS 4304
 QY 600 -----RNIIAAGA---DESVVKEAREVINSSTEGLLNIDKIRKILSGYIVEIDTE 650
 Db 4305 IPKLEINNDIIVGPIKKEIVVKNQK--NNIDVDLGLQINPKIAHLR-FIAKFKSTN 4361
 QY 651 GLKEVIND--RYDMLNIISSL-RQDGKTFIDF-----KKY----- 681
 Db 4362 -----NDIETNVINGSSLVNDGKTSIRFTLNNLKANKLYSLVDVYLVNNSNTIVE 4415
 QY 682 NDKLPLYSNPKYKVNVAVTKENTIIINPSENGDTS 717
 Db 4416 SNKLP-KLNNINYQIK---INKSHTII--SKNGEWS 4445

RESULT 6
 T28676
 rhoxy protein - plasmodium yoelii (fragment)
 C/Species: Plasmodium yoelii
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
 C/Accession: T28676; A45521
 R/Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
 Mol. Biochem. Parasitol. 76, 329-332, 1996
 A/Title: Comparison of two members of a multigene family coding for high-molecular mass
 A/Reference number: Z20507; MUID:97077455; PMID:8920022
 A/Accession: T28676
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-2401 <SIN>

A;Cross-references: EMBL:U36927; NID:G1041784; PID:G1041785; PIDN:AAB41263.1
R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A;Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple cd
A;Reference number: A45521; MUID:91101660; PMID:2270106
A;Accession: A45521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 2260-2401 <KEE>
A;Cross-references: GB:M34281

Query Match 5.0%; Score 188.5; DB 2; Length 2401;
Best Local Similarity 21.7%; Pred. No. 0.26; Indels 221; Gaps 43;
Matches 182; Conservative 119; Mismatches 316;

QY 1 EVKQENRLNSESSESSQGLLYGFFSDLN-----FQAPMVVTSSTGDLSI 45
DB 402 EVAKENVQNVYKSNILKEIKHYNDQINIDNIKEAKQNYDQKEHKMTIPPENMYQK 461
QY 46 PSSELENIPSNQYFQSAI--WSGFIKY---KSDGYTFATSDADNHVTMWVDDQEV---I 97
DB 462 PSIEIKMKDE---FLSKVKNYNDPDKYKSEHNKFTLTNKTIVSDVEEIKKYE 518
QY 98 NKASNSKIRLEKGRLYQIQYQRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELK 157
DB 519 NKFNDSKSLNETKK--SIESEYQINILTKVDD--YIKVCLNTNELITNCHNQKTLK 573
QY 158 QKSNRK--KRSTAGPTVPRDNDGIPD-----SLEVEGYTVDVKNRTFLSP 205
DB 574 DXLNQNIKTETNSIDIKYTDKFNILTDKKTLETFTGLSLNHNESNKELLTYFYD 633
QY 206 WLSNTHKKGKLYKSSPEKSTASDPYSDPKVTGRIDKNVSPPEARPLVAAPVHVVD 265
DB 634 LKANKKNENMLYQFNEK---EKAVEDIKKQVNDKLVSN-----LEIT 677
QY 266 MENILSKNEDQSTONQSTETISKNSTSTSTHTSEVHGNAEVHASFDDIGGSVAGFS 325
DB 678 IVTSIYNED-----TENE---IGKSIELLNTKYLE-----KYKANVT 713
QY 326 NNSNSTVAT--DHSLSLAGERTWAETMGLNTADTASLNANIRVNTGTAPIVNVLPVTLV 384
DB 714 NLNEIKKLDVDFQFGK-----EKNIKYPDEN--KIKNDIDTLNQK 754
QY 385 LGKN--OTLATIKAD--ENOLSOI-----LAPN---NYYP---SKNLAPIALNAQK-- 425
DB 755 IDKSIETLTETIKGNSNHIDEIKGQIDKLVKVPNTMFNEDPKETEKIENIVEKIDKKK 814
QY 426 -----DFSPPTVMYNOFLEKTKQLRLDTQVYGNIAFYNFENGVRVDTGNSWSE 479
DB 815 NIYKEIDKLLNEISKIENDKTSLEKJNINLSYKSLGNLFLOQIDEEKKAEHTIKAME 874
QY 480 V----LPOIQTETARIIFNGKDLNL-----VERRIAAVNPSDP-----LETT-----K 518
DB 875 AYIDDLNLIKKSCEI---EKENVINMDIKMDIHKEMKALNISHDDYKIYHTTSKNHEEK 931
QY 519 PMTWLKEALKIAGFNEP---NGNLOYQKQDITEFPDFNFQDQTSQNIKNQALAEATNIY 575
DB 932 ISDIRKNSLKIIQDFSEESYINDIKKELEKNVLE-----SONNNTDINOYLSKIE--NIY 984
QY 576 TVLDKIKLNAMILIRDKRF--HYDRNNIAGADESVVKEAHREVINSSTEGLLINIDK 633
DB 985 NIL---KLNKIKIIDKVEYTDIEKN-----KKINAELSNS-----1020
QY 634 DIRKILSGYIVETDEGLKE-----VINDRY---DMNITSSLRQDQKTIIDPKYND 683
DB 1021 --EKI-----ITQKENSLLKCEQSKIKSTIDDNVSECIKNITNL-----KTVIVNEKN-- 1069
QY 684 KLPVLYSN--PNYKNV-----YAVTKENTIINPSENGDTSN--GIKKILIFSCK 730
DB 1070 -INTYFKNAZEYQNVSLNFNEMADTKSQYLINIKKNGTNNYDNIKELKHKKK 1126

RESULT 7

S05603
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (s
N;Alternate names: gp195 surface antigen
C;Species: Plasmodium falciparum
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jun-2000
C;Accession: S05603; S04850
R;Myler, P.J.
submitted to the EMBL Data Library, April 1989
A;Reference number: S05603
A;Accession: S05603
A;Molecule type: mRNA
A;Residues: 1-1639 <MYL>
A;Cross-references: EMBL:X15063; NID:G9896; PIDN:CAA33163.1; PID:G9897
R;Myler, P.J.
Nucleic Acids Res. 17, 5401, 1989
A;Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plas
A;Reference number: S04850; MUID:89345116; PMID:2668887
A;Accession: S04850
A;Molecule type: mRNA
A;Residues: 1504-1639 <MYL2>
A;Cross-references: EMBL:X15063
C;Superfamily: major merozoite surface antigen
C;Keywords: glycoprotein; merozoite; surface antigen
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1639/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 4.9%; Score 185.5; DB 2; Length 1639;
Best Local Similarity 20.5%; Pred. No. 0.21;
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

QY 5 ENRL-INSESSESSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSELENIPSENQYFOSA 63
DB 927 ENTILSGKNYIQELIGQKSE-NF-----YERILKDSSTFTYNE 965
QY 64 IWGFIKVKKSDEYTFATSDADNHVTMWVDDQVINKASNS-----NKIRLEKGRLY- 114
DB 966 SFTNFVKSADD-----INSLNDESKRKLLEDINKKLTQLSPDLYNKYKLERLPD 1020
QY 115 -----QIKIQVORENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNRKKES 168
DB 1021 KKKTGVKQKQIKKLTLLKEQLESKL--NSLANPKHVL--QNFVFFNFKGAETAEIN 1076
QY 169 TSAGPTVPRDNDGIPDSLEVEGYTVDVKNKRTFTSPWISNIHEKKGLTKY---KSSP-- 223
DB 1077 T-----LENTKILLGHY-----KGLVKNYNGESSPLK 1103
QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPPEAR-----HPLVAAPVHVVD 265
DB 1104 TLSEESIQTEDNYASLENFKVLSKLEGKLDNLNLEKKLSYLSGLHLIA-----E 1156
QY 266 MENILSKNEDQSTONQSTETISKNSTSTSTHTSEVHGNAEVHASFDDIGGSVAGFS 325
DB 1157 LKEVIRKNK---YTGNSPSNNP-----DVNNALSEYKFLPEGTDVATVS 1200
QY 326 NNSNSTVAIDHSLSLAGERTWAETMGLNTADTASLNANIRVNTGTAPIVNVLPVTLPTT----- 381
DB 1201 ESGSDTLEQSQPKKPASTHVGAES--NTITTSQ--NVDDDEVDDVLIIVIPGESEEDYDDL 1256
QY 382 -SLVGNKQTLATKADENQISQILAPNYYPSKNLAPIA---LNAQKDFSSPTITWYNN 437
DB 1257 GQVVTGEAVTPSVI---DNILSKI---ENEYEVLYLKLAVGTVRSKKQLENNVMTFNN 1310
QY 438 -----QFLEKTKQLRLDTQV-YGNIA-----YFENGVRVVDTSNNSSEVL 481
DB 1311 VKDILNSRFNKNRKNV-LESDLIPYKDLTSSNYVVKDPYKFLNKEKROKFLSSVNYIK 1369
QY 482 PQIQTETARIIFNGKDLNVERRIAANVPSDPL-----ETTKPDM-TLKEALKIAGF 533
DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSKSDLSIKKYNKQGE 1410
QY 534 NEPN-----GNLOYQKQKITE-----PDFNFQDQTSQNIKNQALAEATNIYT 576
DB 1411 NEKLPFLNNIETLYKTVNDKIDLFVHLEAKVLYTYEK---SNVEVKIKELN--YLKT 1465


```
Qy 577 VLDKIKLNKAKN--ILIRDKFHYDRNN-----IAGV-ADESUVKEAHRVINSSTEGLL 628
Db 1466 IODKLADFKKNNFVGIADLTSDYNNHLLTKFLSTGMVFENLAKTVLSLNDLQGG-M 1524
Qy 629 LNIDKD--IRKILSGYIVEIDTGLKEVINDRDMNLNLSLRQDGTTFIDFKKYNKLP 686
Db 1525 LNISOHQCVKK-----OCFQNGGCFRHLDE-----RECKCLINYQEGDKC- 1566
Qy 687 LYISNPYKNVYAVTKENTTIINPSENG-----DTSNIGIKI 724
Db 1567 --VENPNPTCN-----EN-----NGGCDADAKTEEDSGSNG-KKI 1599

RESULT 8
E90598
membrane nuclease, lipoprotein [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: E90598
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:1133084
A:Accession: E90598
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1125 <CUR>
A:Cross-references: GB:AU445566; PID:gl4090108; PIDN:CAC13866.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV 6930
A:Genetic code: GCG3

Query Match 4.9%; Score 184.5; DB 2; Length 1125;
Best Local Similarity 19.1%; Pred. No. 0.13;
Matches 165; Conservative 162; Mismatches 330; Indels 207; Gaps 40;

Qy 3 KOENRLNESSSQGLGYFDFLNFOAPMVVTSTTGDLSIPSELENT:PSNQYFQS 62
Db 120 ENDNSVNNKENS--NDEN-----LTLTKVRLGHNV-LNQSGNIPTKMLAISK 168

Qy 63 ATWS-----GFIKV-----KASDRTTATSADNHTVWVDDOE-----VI 97
Db 169 VILHNKLDIVGLTEITNENGVKTIINELNKSSDKWYIVSLKKGTTGGSGQBEHVGI 228

Qy 98 NKASNSNKIRLE-----KGRLOIKIQORENPTKGL-----130
Db 229 YK---ENKLTLSGFDDEKSKGFKYENKLV---DDPFKGQKIDFVRPPGVKFKSGNI 282

Qy 131 --DFKLYW--TDSQNKKE---VSSDNLQLPKQKGSNSRKRSTAGTVPDRDNGI 183
Db 283 KNDFTVVFHMSDAPGVKEERGBISAKGYSQGGHVAEALRTKEVMEYFDSI-----DGV 337

Qy 184 PDSLEVEGYT-----VVDVKNKRTFLSP-W--ISNIHEK---K 214
Db 338 NNELFMGDTNFKLGNEAKAPFLLOSQYKSLIKVKNATSLAORWGEYAHYOKIYK 397

Qy 215 GLTKYKSSP--EKWSTASDPYSDFEKVTGRIDK-NVSPPEARHP-----LVAAYPI 261
Db 398 GDLKVENSGFYDLWKVFDNLLNKEBFKTVESRSKKAOKYKGEYSVYLHAISDHTI 457

Qy 262 VHVDMENILSKNEDOSTQNTDSEITISKNTSTSEVHGNAEVHASFIDIGGSVS 321
Db 458 VYTDL--ILQKDD-----QNKSENK--DENNSDKQNDKQDNLSSKNDTPQSKESSPOI 509

Qy 322 AGFSNSNSTVAIDHLSLAGERTWAEITMGLTADTARLNANIRYVNTGTAPIYVLP 381
Db 510 DASQNSNTTNEKDKLDSODE---SKNNAIKSQNDQKDSNLSSKNDTPQSKESSPOI 566

Qy 382 SLVIGKQTLATKADENQLSOLAPNYPYKSLAPIALNAQKDPSSPTITWYNQFLE 441
Db 567 NPNLENNQSHSNGENDSDSKQNTSNRQTKN-----DLRSQKQKNTL-----611
```

```
Qy 442 LEKTKQLDLTDQVYGNATYVNFENGVR---VDTSNKNSEVLPOIQETTARIENGKDL 498
Db 612 ---TKXPSNSNSNVTNKTQNNNSSTKQDEIDTSA-----KTQDSTNSLNKNEBKT 661
Qy 499 NLVVERRI-----AAVNPSPDLETTKPDMLTKEALKI--AFGFNEPNNGNLOYQOKDITE 549
Db 662 NQVEIKTNTESNNSNSTKQENSSTKKEEISKSSESNNVNSSTNKQENIDNKKEEISK 721
Qy 550 FDFNFDQTSQNIKQ-LAELNATNIIYTVLDKILNKNKN-----ILLJEDKFFHYDRNI 603
Db 722 SESNVNNSNSTTQOETPETNESQNNVIGKPNNSQNLNQAIDVSAKKVIGYWNINE 781
Qy 604 AVGADESUVKEAHRVINSSTEGLL-----LNIDDKIRKILSGYIVEIDTGLKEVINDR 659
Db 782 SVGKSASAKAFKAVKVIDHNDKLDLVGIGLVHEETLTKI-----VEEMNKLSDSDK 834
Qy 660 YDMLNLSLRQDGTTF-IDFKYNDKLPYISNPNYKNVAVTKENT---IINSENGD 715
Db 835 W--VQVISEKKQGEFPPVNLARYIG--VIYKEN---KFNIESFKQNTNKGHLNENQPMN 887
Qy 716 TSTNIGIKI-----LIFFSKKG 731
Db 888 SSFNTSEKVSRYVRPPGKIFKSTKG 911

RESULT 9
B64635
toxin-like outer membrane protein HP0922 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: B64635
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpik, P.D.; Smith, H.O.; Fraser, C
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: B64635
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2529 <TOM>
A:Cross-references: GB:AE000602; GB:AE000511; NID:g2314060; PIDN:AAD07969.1; PID:g23140

Query Match 4.9%; Score 184.5; DB 2; Length 2529;
Best Local Similarity 21.0%; Pred. No. 0.44;
Matches 158; Conservative 94; Mismatches 261; Indels 241; Gaps 36;

Qy 123 ENPTEKGLDFKLYWTDSONK---KEVISSDNLQLPKQKGSNSRKRSTAGTVPDR 178
Db 729 QNPPAESVWSGVTL--QNKYNSKGIYID---PNLSQSGSGNLTSTYTANLF--- 779
Qy 179 DNDGIPDSLEVEGYTVVDVKNKRTFLSP---WI-----SNHEKKGLTKYK 220
Db 780 ---GRSFVNIQNGTLIIGNNTSVNSGLIWIGHGFGYITGTFSAANIY----LTNNF 832
Qy 221 SSPEKNKT-----ASD-----PYSDEPKVTGRIDKNVSPPEARHPVAAAYPI 262
Db 833 KTGEGVNSCGGANITFKASDNITMDGLNNAEAVTKMIOQTGAS---QHSYATPDALN 889
Qy 263 HVDMENTILSKNEDQSTQNTDSEITISKNTSTSRHTSEVHGNAEVHASFIDIGGSVSA 322
Db 890 NISVTNSFSF-----DMTWGKFPSPKAKNISFS-----NASF-----S 921
Qy 323 GFSNSNSTVA---IDHLSLAGERTWAEITMGLTADTARLNANIRYVNTGTAPIYVNL- 378
Db 922 GFTNPGSSSVISANATNSLSFINR-----LNGGAVYLNQANSLIFN-NQAVFNVLY 973
Qy 379 -----PTSLVLGKQTLATIKADENQLSOLAPNYPYKSLAPIAL--NAQKDFS 428
Db 974 SRGTSENFATTLQGLGNTNFTLSS-----QSLNENFGDITLQNNANITLGNKSOAFAK 1025
```


Db 485 QTKLENKFTFSLNHEANNELI-KYFSDLKANGLINENMLYNQFTEKEKTFNDIKEX 543
Qy 57 NOYFQSAIWGFIKVKKSDYFATSAADNHTVMWVDD-----QEVINKASNSKIRLEK 110
Db 544 NIHINEISKIEIKIHAS-IYNISEETEREIGINIESLNTKVFVKENVTNLNIK-EK 601
Qy 111 GRLYQIKIQORENPEKGLDFLYWTDQNK-----KEVSSDNLQ----- 152
Db 602 LKHYDF-----SDFGKEGNIKYTDKIKKINDIMAVSQIDOHINGLDDIQKSES 652
Qy 153 -LPELKOKSSNRKGRSTASGTPVDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIH 211
Db 653 YVSEMEQINKLEKVSNT-----EISND-----NVEG-----IKKQII-----VTKID 692
Qy 212 EKKG-----LYKXSSPEKMTASDPYSDPEKVTG-----RIDKNVSPARHPL 255
Db 693 KKKNIYEEINKLSBKIKEMKTSLEKVDNLSYGQNLGNLFQIDEE-KKAENTI 751
Qy 256 VAAPIVHVDMENIILSKEDQSTQNTDSETRTISKNTSTSTHSEVHNAEVH-----A 311
Db 752 KSMEAYID-DLDNKKKSQIE*EMDIKMDINKEMEALKISHDDDKCKDKSKHKNENIS 810
Qy 312 SFDIGGSVAGSFGSNNSTVAIDHSLSLAGERWAEWTGLTADTARLANIRVYWTGT 371
Db 811 DIYKSKIKQDFSR-ESDINDIKNLQ-----KNVSSQHNNSDINOCLNEV 857
Qy 372 APIYVNLPTTSLVLGKQNTLTIKADENQSLAPNNYPS-----KNLAPIALNAQK 425
Db 858 ANIYNILKL-----NKIKIIDKVKVETSEIEKKNKNINDELNNSEK 899
Qy 426 -----DFSSPTITWYNOFLE-----LEKTKQLRLDTD-----QVYGN 458
Db 900 VIKIBGDLSEKCRKINSTLDDKIDECIKNINVLKKNILNEETNTNHPKNAEETNK 959
Qy 459 IATYVNFGRVVDTSNWSSEVLPOIQTETARIIFNGKDLNVLVERRIAAVNSDPLETK 518
Db 960 IVLNSFN-----IEMADNKSOYLEIKKNGT-----NDHDYNIKELK-SHKDKSNGYK-TE 1010
Qy 519 PDWTLKALATAGFNEPNCLOYQKGD-----ITEFDNFDOOTSNIKQALAE 568
Db 1011 ADQNKKAQK-----NKEPQYKEEVTVLLNRYAVEUKNPD-KTKVDSQIITKE 1061
Qy 569 LNATNIVTVLDKIKLNAMN-----ILTRDRFRHYDRNNIIVAGADESVKEAHREVIN- 621
Db 1062 IKAHNYCTLESQSEKKNKNEKIHIEDEVANNDKSKAITSIKVSEPFKTKIKI 1121
Qy 622 ---SSTEGLLNIDKIRKILSGYIYEIEDT-----EGKKEVINDRYDMLNLSLR 669
Db 1122 NEIRTKSDCLKETNDEKQISLSDTQETKLTENGKOLKLTESL-----LESLEK 1172
Qy 670 QDGKTFIDFKKNDKPLIYSN-----PNYKVVYAVTKENTIINPSENGTSTNG 720
Db 1173 KQKKNIEDQKDELDEVNSKIKNTENTVNOCKKGYEIGI--VEKINEIAKTKNQIESTKE 1230
Qy 721 IKK 723
Db 1231 LIK 1233

RESULT 12

T31102
filamentous hemagglutinin 1 - Haemophilus ducreyi
C:Species: Haemophilus ducreyi
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T31102
R:Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.
J. Bacteriol. 180, 6013-6022, 1998
A:Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
A:Reference number: Z20984; MUID:99030326; PMID:9811662
A:Accession: T31102
A:Status: preliminary; translated from GS/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-4152 <WAR>

A:Cross-references: EMBL:AF057695; MID:g3929017; PID:g3929018; PIDN:AAC79757.1
C:Genetics:
A:Gene: lrpA1

Query Match 4.8%; Score 181; DB 2; Length 4152;
Best Local Similarity 22.2%; Pred. No. 1.4; 321; Indels 218; Gaps 45;
Matches 18; Conservative 118; Mismatches

Qy 10 NESESQGLLYYFSDNLNFQ-APMVTSSTG-----LSIPSELENISENQYF 60
Db 162 SRTESTSQ-LVGKLANIQLQKEAKLIILNQVTDGHSNIQGALEAVAGKADLIIVN----- 216
Qy 61 QSAIWGSGFIKVKKSDYFATSAADNHTVMWVDDQVINKASN-----SNKIRLEKGRLY 114
Db 217 PNCITLNGVKTINTDRFVVSST-----DIIPRENGLLSVRGKVTIDKGGVA 264
Qy 115 QIKIOY---ORENPTEKGLDFKLYWTDSONKKEVSSD-----NLQPELKOKKS 161
Db 265 TNGLSHFVVARNIQDKG---KITVAKTENQKSVNPANITFAAGSLNLTNLTREATFISS 321
Qy 162 NSRKKESTSAGTVPDRNDGIPDSLEVEGYTVD---VKNKRTPLSPWISNIHEKKGLT 217
Db 322 GTSR---TSDTFAI-SAGSAGMYGSNIKFFVTDKAGVKKKGIIFSENDINIKMDGNA 377
Qy 218 KYKSSPEKSTASDPYSDPEKVTGTRIDKNVSEARHPLVAAPIVHVDMENIILSKEDQ 277
Db 378 SLKELYAK-----KQIDILAKDIELTEKGOLOANNKIILNSTGKINLRNASEV 425
Qy 278 STQNTDSETRTIS-KNTSTSRHTSEVHNAEVAHSPFDIGGSVAGSFGSNNSTVAIDH 336
Db 426 SADNVNKSLENALENASMSANSLDVIYTKLEVRS-----SKVSAGTANIKASNIILDG 480
Qy 337 SLSLAGERTW-----AETMGLNTADTARLN-----AMIRYV---T 369
Db 481 SSVANKITLNTVNTNATLNQSKLSAKDMELNVTHITLNTTSKLSAQKANIKNITENLTN 540
Qy 370 GTAPIVNLPTTSLVLGK-----NOTLA-----TIKADENO-LSOILAPNN-- 409
Db 541 GEA-----SLVAEKLDINAIDKIYNNGTIAGLTANITTKALENEDNALILAHQNLN 591
Qy 410 -----YY-----PSKNLAPIALNAQKDFSS-----TPITMNVNQPLEKTKQL 448
Db 592 FTVNGSHYVNGKDIVSKCKAIVTFSSNDSFTSNGSKLYDAQNNLTNVVNNFNITQGSBII 651
Qy 449 RLDTDOVYGNI---ATYAPEN-GRVVDTSNWSSEVLPOIQTETARIIFNGKDLNVLVER- 503
Db 652 -----LHGNTVLNAKGNFTNSGNLTTKELNINIESFI---NAGNLTGKNLEVHSNT 702
Qy 504 -----RIAAVNSPDLPTTKEDMTL-KEALKIARG-----FNEPKNLOYOQKDIETEF 550
Db 703 TVKNQDKLYSIEENLISKSTDFTNNGTLLGLEALKIASGNGFTNASNGSLA-SNKSLDIY 761
Qy 551 DENF-DQOQTSQNIKNQALAEALNATNIYTVLDK--IKLNKKNILIRDKRFHYDRNNIIVAG 607
Db 762 GNNFTNGTIESVKS-----LNTNNYTFINNATIKSYGVLTISQO-NFTNDSNGTWSH 816
Qy 608 D-ESVVKEAHREVINSSST-----EGULL-----NIDKIRKILSGYIYEIEDTEGLKEVIN 658
Db 817 DLINITSQAN--IINKLLAGGGGLNLTAKGNITNDSNSTAIAVLHNSNDIN-----LNA 869
Qy 659 RYDMLNLSL-RQDGKTFIDFKKNDKPLIYSNPNYKVVYAVTKENTIINPSENGDTS 717
Db 870 NKKYVNIIGIYQAGNISVEAKLLHNDVKL-----SGNITTTK-----SGNATVK 915
Qy 718 TNGI 721
Db 916 TNSI 919

RESULT 13

S68218
botulinum neurotoxin type A, nontoxic/nonhemagglutinating - Clostridium botulinum (stra
C:Species: Clostridium botulinum

A:Variety: strain NIH
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 02-Jun-2003
C:Accession: S68218; S74301
R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376. 41-44. 1995
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin components
A:Reference number: S67988; MUID:96096783; PMID:8521962
A:Accession: S68218
A:Molecule type: DNA
A:Residues: 1-1193 <FUJ>
A:Cross-references: EMBL:D67030; NID:g2160224; PIDN:BAAL1050.1; PID:d1011710; PID:g11324
A:Experimental source: strain NIH
A:Accession: S74301
A:Molecule type: protein
A:Residues: 1-13,145-155 <FUI>
A:Experimental source: strain NIH
C:Genetics:
A:Gene: ant
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin
F:1-1193/Product: botulinum neurotoxin type A, nontoxic/nonhemagglutinating #status expected

Query Match 4.8%; Score 180; DB 2; Length 1193;
Best Local Similarity 19.4%; Pred. No. 0.24;
Matches 181; Conservative 131; Mismatches 319; Indels 294; Gaps 46;

Qy 42 DLSIP---SSELENIPSENYFQSAIWSGFIKVK--KSDBYTPATS-----ADNHYT 88
Db 241 DLVPIRLSRSELENIYSQNLNIVDLVSGGIDPKFINTDPYFTDNYFSAKVPEDHRN 300
Qy 89 MWVDQEVINKASNSKIRLEK-----GRLYQIKIQY-----QRENTEKGLDF 132
Db 301 IYETEIEGNAIGNDIKLRLQKQFRININDIWEINLNYFSKPSIMWPRFNALKHFYR 360
Qy 133 KLYW-----TDSQNKKEVISS----- 148
Db 361 QYKYKIDYPENYSINGFVNGQINAOQLSLDRNQDIINKPEIINLLNGNVSLMRSNIYV 420
Qy 149 -----DNQLP-----ELKQKSSNRKSTSGAGTPVDRDNDGIDPSLEVEGY 192
Db 421 DGLKSTVDDFVSNKIPINRAYEYHFNNNSDLSLDNNGI-----DNIFEIDVNPY 474
Qy 193 TVD-----VKNKRTFLS--PWISNIHEKGLTKYKSSPKWSTASDPYSDFEKYTG 241
Db 475 KENCDFKSPVQKITSTREINTNIPWPNY-----LQAQNTNNEKFSLS-----SDFEVEVS 525
Qy 242 RIDKQVSPPEARHPLVAAPYVHVDMENII--LSKNEDQSTQNTDSE-----TRTISKNTST 295
Db 526 SKDKS-----LVYSFLSNVMFYLSIKDINSPIDTDKYYLWLREIFRNYGF 571
Qy 296 SRTHSTSEVHGAEVH-----ASFDDIGGSVSAFGNSNS--STVAIDHSLSLAGERT 345
Db 572 DITATQETNCGINKVVTWFGKALNINTSD3FVEEFQNLGAISLINKKENLSMPLIES 631
Qy 346 WA---ETWGLNTADTARLANIRVNTG--TAPIYVNLPT-----TSVLGKQNT 390
Db 632 YEIPNDMLGLPLNDLNEKLFNIYSKNTAYPKKYVYNFLDQWWTQYYSQYFDDICWARRSV 691
Qy 391 LA---TIKADENQLSOLAPNYPYKXKLAPIAL---NAQKPFSS--TPITMN-----435
Db 692 LAQETLIKRIIQKLSVLIIGNSSDNLALMNLTTNTLRDISNESQIAGANVDSFLNN 751
Qy 436 -----NQFLELEKTKQLRLDQVYG--NIATYVNFENGRVRVDTCGWSWEVLPO 483
Db 752 AAICVFESNIYPKFISF-----NEQCINNINIKTEF-----IQKCTNINEDKQL 798
Qy 484 IOETTARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKEAL-----KIAFGFNEPNGN 539
Db 799 INQN-----VFNSLDPEFLN-----IQNKSLFSSETALLIKEETWPELVLYAFKEPGNN 849
Qy 540 -----LOYOGKDI--TEPDFNF-----QOTSQNIKNQELANATYIVT-----577
Db 850 VIGDASGKNTSIEY-SKDIGVYGINSDALVNGSNQSI5FSNDFENGITNSFSIYFWL 908

Qy 578 --LDKIKLNAKM-----NIIIRD---KRPFY--- 598
Db 909 RNLGKDTIKSLGSKEDNCGWEIFQDTGLVFNWIDSGNKGKNIYLSDVSNNSWHYITI 968
Qy 599 --DR--NNIYAGADESVV--KEAHRVINSSTEGILLNIDKDIRKILSGYVIEIEDTEGLK 653
Db 969 SVDRLEKQULLIFDDNLVANESIKELIILNYSNIISLSENNPSYIEGUTILNKPTTS-Q 1027
Qy 654 EVINDRYDMNLINSSLPQDKTIFDKKYNCKPLPLYSNPYKVN---VYAVTKENTI--- 707
Db 1028 EVLSNYFEVLNNSYIRDSNEERLE---YNTKYOLY---NYVFSKPICEVKQNNIYLT 1080
Qy 708 INPSENGDSTNGIKKILIFSKGY 732
Db 1081 INNTNLLNQASKFKLLSINPNKQY 1105

RESULT 14
A:86827
hypothetical protein yqfG [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: A86827
R:Boilotin, A.; Wincker, P.; Mager, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrl
Genome Res. 11. 731-753. 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s.
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: A86827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1072 <STO>
A:Cross-references: GB:AA005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yqfG

Query Match 4.8%; Score 179.5; DB 2; Length 1072;
Best Local Similarity 19.4%; Pred. No. 0.22;
Matches 160; Conservative 142; Mismatches 342; Indels 179; Gaps 30;

Qy 2 VKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGLSLSPSELEN----- 52
Db 300 VEQVDVASSESTQDANSALYPISEASSVTDNLANSISLSDSSISSQTSQSGASSTA 359
Qy 53 --IPSENYFQSA---IWSGFIKVKSDYETFATSADNHVTMWYDDQEVINKASNSKI 106
Db 360 EISYDSENSLSLSSNQINSNSKQSNQSLGSSMSSESEHSNSNSNINETNNSSEI 419
Qy 107 R-----LEKRLYQIKIQYQRENPTKGLDFKL-----YWTDQNKKEVISSDNQLQPELK 157
Db 420 TNILPPSNPTESNSVSDQTSSEASTNSNSISLSLSPSNISSTSDSESATNSDSFNVAEVA 479
Qy 158 QKSSNRKRRSTAGTPVDRDNDGI---PDSLEVEGYTVDKNKRTPLSWNIHEK 213
Db 480 NNSLASVNNSSSVLSSTSTADNLGINSQSNLTKD---SSEISTSGAFLS---SNQTS 534
Qy 214 KGLTKYS---SPEKWS-----TASDPYSDFEKYTGRIKDVSPPEARHPLVAAPYI 261
Db 535 EASTNSNSISLSLSPSNISSTVLESTSSNFSNVAEVANNLSASVNNSSSVLSSTSTA 594
Qy 262 VHYDM-----ENI-----ILSKNEQSTQNTDSETRTISKNTSTSTRTSTSEVH 304
Db 595 DNLEIQFGSDNLTKDSSSEISTSGAFLSNQTSASSNSMSINSPLSLSLTNSESA 654
Qy 305 GNAEVHASFPDIGGSVSAFGNSNSSTVAIDHSLSL--AGERTWAEWTGLNTADTARLANA 362
Db 655 TNQ-----SNSSEATKVDNNSST---HSSNILLNSGNSDSDSDSDSSNL-- 699
Qy 363 NIRVNTGTAPIYVNLPTTSLVGNQTLATIKADENQLSQ-----ILAPNNYPSKCLAP 418
Db 700 -----SSSPNLETNQIISRPSEVNNISENPKVSSNSVQENSTIDHE 742

```
QY 419 IALNAQDFSTPTMNYNQFLELEKTKQLRLDQVYGNATVN---FENGVRVDTGS 475
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 743 MSTNPKSSISPISTSSQOKESQN---LLNTEGINNPIFNSSSSNSAASILT-- 797
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 476 NWSEVLFOIQBTARIIF---NGKDL-----NLVERRIAAVNPSPLETTKPDMT 522
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 798 SYSNNSESSETGLYISNEAQRDNGEISHSPSSNSNNVSSIQSQAILESSKSTN 857
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 523 LK-EALKIAFGFNP-NGNLQYQKQDITEFPNFDQQTSONIKNQLAELNATNIYTVLDK 580
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 858 KRSSLSIINTSHPONEDNOSNSD-----EVKSNVNSVILGOLNSISNKTWN 909
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 581 I---KLNKAKMILIRDKFRHYDRNNAIAGADESVVKEAHRVINSSTEGLLNIDKIRK 637
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 910 LTSQKLSVIYTLPSKSKVTNEKNNSNTVSEKLIKTPQKN-DESQNLGOITALDLFNNK 968
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 638 ILSGYVIEIEDTEGLKEVIN-----DRYDMLNISLROD-KTFTIDFKYNDKLPY 688
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 969 -----EVETMEDSKTVPDKVLNENGRSQNNKTSTIAKDKNKFVKSGSEFNKIL-- 1019
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 689 ISNPYKVNVAVTKEITIINPSENGDTSITNGIKKILIFSKG 731
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1020 -----DSNNILKKTVLKKG 1036
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
A:Accession: A10452
hemolysin [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
R:Accession: A10452
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: A10452
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1635 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC93189.1; PID:GI5981638; GSPDB:GN00175
C:Genetics:
A:Gene: YPO3721
```

```
Query Match 4.7%; Score 179; DB 2; Length 1635;
Best Local Similarity 19.4%; Pred. No. 0.44;
Matches 176; Conservative 137; Mismatches 315; Indels 278; Gaps 44;

QY 8 LNESESSQGLGYFSDLNFOAPMVTSSVTGDLSPSELENIPISENQYFQSAIWSG 67
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 510 LTAENKITTOGIKASAGDVVIDA-----NDVKIGVQKTSNQETTDGKHENLGLG 560
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 FIKYKSDVYTFAT-----SADNHTVMWVDDQEVIN-----KASNSKIRLEKRLYQIKI 118
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 561 GVDHNNNDKYAETSHSSEITADGNILISVDVAITGSKVKATKDFGVOAKEG---GIKI 617
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 QYQRENPTK-----GLDFKLYWTDSONK-----KEVISSDNLOLPELKOKSSNS 163
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 618 DNASTITTNKVDERTGAFDI--TGSSKKANNESEKSTGEVSEVSEANLKI--ISKDVVDV 673
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 RKXRSTSAG-----PTVPD-----RDNDGIPDS-----LEVEGYTV-DVKNKRTFLSPWISN 209
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 674 IGSIVKSAGELGIETILGDINVAQAQEKQIDEQKTLTIDGFTSDGKQKQ----- 723
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 210 IHEKKGL-TKYKSPKWTASDPYSDPEKVTGRI--DKNVSPEARHPVAAVPIVHDM 266
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 724 --YQAGLKVEHTESEKTEKVINHGSLTEGGTVKLEADKVDVFTGSG-LNTTKGDADITA 780
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 ENITLSKNEQOSTQNTD-----SETRISKNTSTSRHTS 301
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 781 ENVSFVAQDTTISNKEKTVGVNAHYTGGMDKAGSAGVSYEETKTDSEKSTAVVSHT- 839
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: May 3, 2004, 19:41:16

Job time : 16.0855 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 8.2539 Seconds

(without alignments)
4636.784 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENRLNSESSESSQGL.....TSTNGIKLILFSKKGVEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3758	99.6	764	1 PAG_BACAN	P13423 bacillus an
2	235.5	6.2	204	1 YPBI_BACAN	P13422 bacillus an
3	185.5	4.9	1630	1 MSP1_PLAFK	P04932 plasmodium
4	185.5	4.9	1639	1 MSP1_PLAFW	P04933 plasmodium
5	174	4.6	1658	1 YM67_YEAST	Q03661 saccharomyc
6	169.5	4.5	1230	1 SMC3_YEAST	P47037 saccharomyc
7	166	4.4	1000	1 W152_YEAST	P43612 saccharomyc
8	165.5	4.4	2334	1 WAPA_BACSU	Q07833 bacillus eu
9	164	4.3	1803	1 YUL3_YEAST	P47024 saccharomyc
10	162.5	4.3	2116	1 MYS2_DICDI	P08799 dictyosteli
11	162	4.3	1957	1 SPOF_SCHPO	Q10411 schizosacch
12	161.5	4.3	979	1 P115_MYCHR	P41508 mycoplasma
13	160	4.2	1276	1 BXD_CLOBO	P19321 clostridium
14	159.5	4.2	2869	1 RBPI_PLAVB	Q00798 plasmodium
15	159	4.2	1385	1 PAT1_SCHPO	O13735 schizosacch
16	158.5	4.2	1167	1 CAGA_HELPU	Q921t1 helicobacte
17	157	4.2	2867	1 RBP2_PLAVB	Q00799 plasmodium
18	155.5	4.1	1024	1 RIP3_MOUSE	P97434 mus musculu
19	155	4.1	1790	1 USOI_YEAST	P25386 saccharomyc
20	154.5	4.1	1208	1 PCPI_SCHPO	Q92351 schizosacch
21	154.5	4.1	1460	1 N159_YEAST	P40477 saccharomyc
22	154.5	4.1	1487	1 MDS3_YEAST	P53094 saccharomyc
23	154	4.1	1037	1 KCC4_YEAST	P25389 saccharomyc
24	153.5	4.1	1233	1 YF16_YEAST	P43597 saccharomyc
25	153	4.1	1029	1 RIP3_RAT	Q9ere6 rattus norv
26	152.5	4.0	1116	1 YK54_AQUAE	O67838 aquifex aeo
27	152	4.0	1420	1 SRB9_YEAST	P38931 saccharomyc
28	150	4.0	918	1 YMJ6_CAEEL	P34487 caenorhabdi
29	150	4.0	1882	1 Y468_MYCPN	P75109 mycoplasma
30	150	4.0	1928	1 MYS1_YEAST	P08964 saccharomyc
31	149.5	4.0	678	1 INCY_YEAST	P53968 saccharomyc
32	149	3.9	1250	1 BXE_CLOBO	Q00496 clostridium
33	149	3.9	1577	1 HLVA_PROMI	P16466 proteus mir

ALIGNMENTS

RESULT 1

ID	PAG_BACAN	STANDARD	PRT	764 AA.
AC	P13423: Q9F5R7: Q9KH69: Q9ROU2;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].			
GN	PAGA OR PAG OR PXOI-110.			
OS	Bacillus anthracis.			
OC	Plasmid PXOI.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1392;			
FN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=69172073; PubMed=3148491;			
RA	Welkos S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppla S.H., Schmidt J.J.;			
RA	"Sequence and analysis of the DNA encoding protective antigen of Bacillus anthracis.";			
RL	Gene 69:287-300(1988).			
FN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=28_33, BA1024, and BA1035;			
RX	MEDLINE=99214082; PubMed=10197996;			
RA	Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;			
RT	"Genetic diversity in the protective antigen gene of Bacillus anthracis.";			
RL	J. Bacteriol. 181:2358-2362(1999).			
FN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=V770-NP1-R / ATCC 14185;			
RX	MEDLINE=20353947; PubMed=10899854;			
RA	Cohen S., Mendelson I., Alboum Z., Kobiler D., Elhanany E., Bino T., Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M., Kronman C., Velan B., Shaffer A.;			
RA	"Attenuated nontoxinogenic and nonencapsulated recombinant Bacillus anthracis spore vaccines protect against anthrax.";			
RL	Infect. Immun. 68:4549-4558(2000).			
FN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sterne;			
RX	MEDLINE=99445483; PubMed=10515943;			
RA	Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K., Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D., Martinez Y., Ricks D., Svensson R., Jackson P.J.;			
RA	"Sequence and organization of pXOI, the large Bacillus anthracis plasmid harboring the Anthrax toxin genes.";			
RL	J. Bacteriol. 181:6509-6515(1999).			
FN	[5]			
RP	DOMAINS.			
RX	MEDLINE=91332080; PubMed=1651334;			
RA	Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;			
RT	"The carboxyl-terminal end of protective antigen is required for receptor binding and anthrax toxin activity.";			

Q96173 homo sapien
P40468 saccharomyc
Q06366 clostridium
P47551 mycoplasma
P11978 saccharomyc
P50495 plasmodium
Q9zkW5 helicobacte
P18494 saccharomyc
Q9c520 lactococcus
Q03213 saccharomyc
Q53654 staphylococ
P58302 thermoplasm

RL J. Biol. Chem. 266:15493-15497(1991).
 [6] CHARACTERIZATION.
 RX STRAIN=Sterne;
 RC MEDLINE=94327640; PubMed=8051159;
 RA Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;
 RT "Anthrax protective antigen forms oligomers during intoxication of
 RL mammalian cells";
 RN J. Biol. Chem. 269:20607-20612(1994).
 [7] CHARACTERIZATION.
 RX MEDLINE=2112592; PubMed=1207581;
 RC Beaugreard K.E., Collier R.J., Swanson J.A.;
 RA "Proteolytic activation of receptor-bound anthrax protective antigen
 RT on macrophages promotes its internalization";
 RL Cell. Microbiol. 2:251-258(2000).
 [8] TOXIN REGULATION.
 RX STRAIN=Weybridge;
 RC Koehler T.M., Dai Z., Kaufman-Yarbray M.;
 RA "Regulation of the Bacillus anthracis protective antigen gene: CO2 and
 RT a trans-acting element activate transcription from one of two
 RL promoters";
 RN J. Bacteriol. 176:586-595(1994).
 [9] MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.
 RX STRAIN=Sterne;
 RC MEDLINE=95050722; PubMed=7961869;
 RA Singh Y., Klimpel K.R., Arora N., Sharma M., Leppia S.H.;
 RT "The chymotrypsin-sensitive site, FFD315, in anthrax toxin protective
 RL antigen is required for translocation of lethal factor";
 RN J. Biol. Chem. 269:29039-29046(1994).
 [10] MUTAGENESIS OF DOMAIN 4 LOOPS.
 RX STRAIN=Sterne;
 RC MEDLINE=99185012; PubMed=10085028;
 RA Varughese M., Teixeira A.V., Liu S., Leppia S.H.;
 RT "Identification of a receptor-binding region within domain 4 of the
 RL protective antigen component of anthrax toxin";
 RN Infect. Immun. 67:1860-1865(1999).
 [11] MUTAGENESIS OF TRP-375, MET-379 AND LEU-381.
 RX STRAIN=Sterne;
 RC MEDLINE=2102804; PubMed=11178978;
 RA Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;
 RT "Trp 346 and Leu 352 residues in protective antigen are required for
 RL the expression of anthrax lethal toxin activity";
 RN Biochem. Biophys. Res. Commun. 281:186-192(2001).
 [12] MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.
 RX STRAIN=Sterne;
 RC MEDLINE=21438996; PubMed=11554763;
 RA Ahuja N., Kumar P., Bhatnagar R.;
 RT "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are
 RL required for oligomerization of anthrax protective antigen";
 RN Biochem. Biophys. Res. Commun. 287:542-549(2001).
 [13] MUTAGENESIS OF PRO-289.
 RX STRAIN=Sterne;
 RC MEDLINE=21255689; PubMed=11356563;
 RA Khanna H., Chopra A.P., Arora N., Chaudhry A., Singh Y.;
 RT "Role of residues constituting the 2beta1 strand of domain II in the
 RL biological activity of anthrax protective antigen";
 RN FEMS Microbiol. Lett. 199:27-31(2001).
 [14] MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.
 RX MEDLINE=21125576; PubMed=11222612;
 RA Mogridge J., Mouriz M., Collier R.J.;
 RT "Involvement of domain 3 in oligomerization by the protective antigen
 RL moiety of anthrax toxin";
 RN J. Bacteriol. 183:2111-2116(2001).
 [15]

RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.
 RX MEDLINE=21269403; PubMed=11113126;
 RA Sellman B.R., Nassi S., Collier R.J.;
 RT "Point mutations in anthrax protective antigen that block
 RL translocation";
 RN J. Biol. Chem. 276:8371-8376(2001).
 [16] MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;
 RX ILE-239; TRP-255 AND PHE-265.
 RC STRAIN=Sterne;
 RA MEDLINE=22112896; PubMed=12117959;
 RA Chauhan V., Bhatnagar R.;
 RT "Identification of amino acid residues of anthrax protective antigen
 RL involved in binding with lethal factor";
 RN Infect. Immun. 70:4477-4484(2002).
 [17] X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RX MEDLINE=97192099; PubMed=9039918;
 RA Petosa C., Collier R.J., Klimpel K.R., Leppia S.H., Liddington R.C.;
 RT "Crystal structure of the anthrax toxin protective antigen";
 RN Nature 385:833-838(1997).
 [18] REVIEW.
 RX MEDLINE=21428689; PubMed=11544370;
 RA Mock M., Pouet A.;
 RT "Anthrax";
 RN Annu. Rev. Microbiol. 55:647-671(2001).
 CC -!- FUNCTION: One of the three proteins composing the anthrax toxin,
 CC the agent which infects many mammalian species and that may cause
 CC death. PA binds to a receptor (ATR) in sensitive eukaryotic
 CC cells, thereby facilitating the translocation of the enzymatic
 CC toxin components, edema factor and lethal factor, across the
 CC target cell membrane. PA associated with LF causes death when
 CC injected, PA associated with EF produces edema. PA induces
 CC immunity to infection with anthrax.
 CC -!- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a
 CC protective antigen (PA), a lethal factor (LF) and an edema factor
 CC (EF). None of these is toxic by itself. PA+LF forms the lethal
 CC toxin (LeTx); PA+EF forms the edema toxin (EdTx). PA-63 forms
 CC heptamers and this oligomerization is required for LF or EF
 CC binding. Once activated, at low pH, the heptamer undergoes
 CC conformational changes and converts from prepore to pore inserted
 CC in the membrane, forming cation-selective channels.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- DOMAIN: The molecule is folded into four functional domains. Each
 CC domain is required for a particular step in the toxicity process.
 CC Domain 1 contains two calcium ions and the proteolytic activation
 CC site. Cleavage of the PA monomer releases the subdomain 1a, which
 CC is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is
 CC part of the remaining 63-kDa fragment (PA63) and contains the
 CC binding sites for LF and EF. Domain 2 is a beta-barrel core
 CC containing a large flexible loop that has been implicated in
 CC membrane insertion and pore formation. There is a chymotrypsin
 CC cleavage site in this loop that is required for toxicity. Domain 3
 CC has a hydrophobic patch thought to be involved in protein-protein
 CC interactions. Domain 4 appears to be a separate domain and shows
 CC limited contact with the other three domains: it would swing out
 CC of the way during membrane insertion. It is required for binding
 CC to the receptor; the small loop is involved in receptor
 CC recognition.
 CC -!- PTM: Proteolytic activation by furin or a furin-like protease
 CC cleaves the protein in two parts, PA-20 and PA-63; the latter is
 CC the mature protein. The cleavage occurs at the cell surface and
 CC probably in the serum of infected animals as well; both native and
 CC cleaved PA are able to bind to the cell receptor. The release of
 CC PA20 from the remaining receptor-bound PA63 exposes the binding
 CC site for EF and LF, and promotes oligomerization and
 CC internalization of the protein.
 CC -!- MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were
 CC performed that showed that the residues present in the small loop
 CC of domain 4, and not the ones in the large loop, are involved in
 CC receptor recognition.
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.


```
RX MEDLINE=89172073; PubMed=3148491;
RA Welkos S.J., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppa S.H.,
RA Schmidt J.J.;
RT "Sequence and analysis of the DNA encoding protective antigen of
RT Bacillus anthracis."
RL Gene 69:287-300(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sterner;
RX MEDLINE=99445483; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
RA Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
RA Martineez Y., Ricke D., Svensson R., Jackson P.J.;
RT "Sequence and organization of pXOI, the large Bacillus anthracis
RT plasmid harboring the Anthrax toxin genes."
RL J. Bacteriol. 181:6509-6515(1999).
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22589; AAA22636.1; ..
CC EMBL; AF065404; AAD32415.1; ..
CC PIR; G59104; G59104.
CC PIR; I39933; I39933.
CC HSP; P3423; IACC.
CC Hypothetical protein; Plasmid; Transmembrane.
CC TRANSMEM 162 182
CC CONFLICT 93 93 T -> A (IN REF. 1).
CC CONFLICT 184 204 KSCNCLLYVEVSQLMNSVYF -> NHVIVYLSM (IN
CC REF. 1).
CC SEQUENCE 204 AA; 23029 MW; E1657B23AB4273FD CRC64;
Query Match 6.2%; Score 235.5; DB 1; Length 204;
Best Local Similarity 34.6%; Pred. No. 2.6e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;
QY 587 MNTILIRDKRFHYDRNNIAGVADSVVKEAHREVINSSTEGLLINIDKIRKILSGYIVEI 646
Db 1 MNTILIRDP-VHYDNGNIYGVDDSYLNKAYKQLNWSGVSGLNDEVDYQALSGYMLQI 59
QY 647 EDTE-----GLKEVINDRYDMLNSSLRQDKTFFDKYNDKPLYSNPN 693
Db 60 KPSPNHLTNSPVITITLAGKDSGVGELYRVLS-----DGTGFLDNKFDENRSLV-DPG 112
QY 694 YKQNVYATKEN-TIINPSENGDTSTNGIKKILIFSKGVEI 734
Db 113 DDVVVYAVTKEDFNAVTRDENGNI-A-NKLNVLVLSGKIKEI 153
RESULT 3
MSPI_PLAFK
ID MSPI_PLAFK STANDARD; PRT; 1630 AA.
AC P04932;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
RA Stunnenberg H., Bujard H.;
```

```
QY 1 EVKQENLLNESBSSQGLGYPSDLNFOQPMVVTSTTGDLSIPSELENIPSENOYF 60
Db 30 EVKQENLLNESBSSQGLGYPSDLNFOQPMVVTSTTGDLSIPSELENIPSENOYF 89
QY 61 QSAIWSGFIKVKSDXYTFATSDAHVMTWYDDQEVINKASNSKIRLEKRLYQIKIY 120
Db 90 QSAIWSGFIKVKSDXYTFATSDAHVMTWYDDQEVINKASNSKIRLEKRLYQIKIY 149
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQELPELKQKSNRSKSTSGAPVPRDN 180
Db 150 QRENPTKGLDFKLYWTDSONKKEVSSDNLQELPELKQKSNRSKSTSGAPVPRDN 209
QY 181 DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKVT 240
Db 210 DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKVT 269
QY 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKVTSTHT 300
Db 270 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKVTSTHT 329
QY 301 SEVHGNAEVHASFDDIGSVSAGFNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
Db 330 SEVHGNAEVHASFDDIGSVSAGFNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 389
QY 361 NANIRYVNTGTAIYVNLPTSLVLGKQNTLTIKADENQQLSQILAPNNYPSKQLAPIA 420
Db 390 NANIRYVNTGTAIYVNLPTSLVLGKQNTLTIKADENQQLSQILAPNNYPSKQLAPIA 449
QY 421 LNAQKDFSTPTIMYNOFLEKTKQLRLDQVYGNIAVYNGVRVDTGNSNVEV 480
Db 450 LNAQKDFSTPTIMYNOFLEKTKQLRLDQVYGNIAVYNGVRVDTGNSNVEV 509
QY 481 LPQIOETTARIIFNGKDLNVERIAAVNPSDPLETTPDMTLKALKIAFGFNEPNGNL 540
Db 510 LPQIOETTARIIFNGKDLNVERIAAVNPSDPLETTPDMTLKALKIAFGFNEPNGNL 569
QY 541 QYQCKDITEFPDFNFOQTSQNTKQALAEANVITVLDKIKLNKMLIRDKFHYDR 600
Db 570 QYQCKDITEFPDFNFOQTSQNTKQALAEANVITVLDKIKLNKMLIRDKFHYDR 629
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLINIDKIRKILSGYIVEIDTEGLKEVINDRY 660
Db 630 NNIAVGADESVMKEAHREVINSSTEGLLINIDKIRKILSGYIVEIDTEGLKEVINDRY 689
QY 661 DMLNSSLRQDKTFFDKYNDKPLYSNPNYKQNVYATKENTIIINPSENGDTSTNG 720
Db 690 DMLNSSLRQDKTFFDKYNDKPLYSNPNYKQNVYATKENTIIINPSENGDTSTNG 749
QY 721 IKKILIFSKKGVEIG 735
Db 750 IKKILIFSKKGVEIG 764
RESULT 2
YPB1_BACAN
ID YPB1_BACAN STANDARD; PRT; 204 AA.
AC P13422; Q9X377.
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein pXOI-111.
GN pXOI-111.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
```

RT "Polymorphism of the precursor for the major surface antigens of
 RL Plasmodium falciparum merozoites: studies at the genetic level.";
 ENMO J. 4:3823-3829(1985).
 RN [2]
 RP REVISIONS, SEQUENCE FROM N.A.
 RA Pan W., Tolle R., Bujard H.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42
 CC kDa and 19 kDa antigens which are the major surface antigens of
 CC merozoites. The maturation take place during schizont.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X03371; CAA27070.1; -;
 DR InterPro: IPR006209; EGF_like.
 DR Pfam: PF00008; EGF; 1; _like.
 KW Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;
 KW Transmembrane; GPI-anchor.
 FT SIGNAL 1 19
 FT CHAIN 20 1630
 FT DOMAIN 67 84
 FT TRANSMEM 1614 1630
 FT CARBOHYD 97 97
 FT CARBOHYD 259 259
 FT CARBOHYD 755 755
 FT CARBOHYD 759 759
 FT CARBOHYD 774 774
 FT CARBOHYD 835 835
 FT CARBOHYD 911 911
 FT CARBOHYD 955 955
 FT CARBOHYD 1049 1049
 FT CARBOHYD 1156 1156
 FT CARBOHYD 1165 1165
 FT CARBOHYD 1436 1436
 FT CARBOHYD 1517 1517
 SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;
 Query Match 4.98; Score 185.5; DB 1; Length 1630;
 Best Local Similarity 20.5%; Pred. No. 0.12;
 Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;
 QY 5 ENRL-LNESSSSQGLLYGYFSDINFAQPMVVTSTTGDLSPSSSELENIPSENYQFQSA 63
 DB 918 ENLISLGNKNIYQELIGQKSSSE-NF-----YKILKDSDTFYNE 956
 QY 64 IWSGFIKVKSDYTPATSDADNHVTWVDQEVINKASN-----NKIRLEKGRLY- 114
 DB 957 SFTNFVKSQADD-----INSLNDSKRGKLEEDINKLTKLQSLFPLYNKYLKLERLFD 1011
 QY 115 -----QIKIQYQENPTKGLDFKLYWTDSONKKEVISSDNLQPLKQKSSNSRKRS 168
 DB 1012 KKKTVGVYQIKLTLKQLESKL--NSLNKPKVYL--QNSVFPNKKEAEIAETEN 1067
 QY 169 TSAGTVPDRNDGIPDSLEVEGVTVVKNKRTPLSPWISNIHEKKGLTKY----KSP-- 223
 DB 1068 T-----LENTKILKHY-----KGLVYKYNNGESSPLK 1094
 QY 224 ---EKWSTASDPYDFE-----KVTGRIDQNVSPPEAR-----HPLVAAYPIVHVD 265
 DB 1095 TLSEESTQDENYASLENFKVLKLEGLKDNLNLEKKLSYLSGLHLHIA-----E 1147
 QY 266 MENILSKNEDQSTQNTDSETRTISKNTTSRTHTSEVHGNAEVHASFPIGGSVSAQGS 325
 DB 1148 LKEVIKKNK---YTGNSPSENNT-----DVNNALESYKKFLPEGTDTVATWS 1191

QY 326 NNSSTVAIDHSLSLAGERTWAETWMTGLNTPADTARLNANIRYNTGTAPIYVLPPT---- 381
 DB 1192 ESGSDTLEQSQPKKPASTHGAES---NTITTSQ-NVDDEVDDVIIVPIFGSEEDYDDL 1247
 QY 382 -SLVLGNQTLATIKADENQLSQILAPNPNYPSKNLAPIA---LNAQKDFSPTPTMKN 437
 DB 1248 GQVVTGEAVTPSVI---DNILSKI---ENEYEVLYLKLPLAGVYRSKLSKLENNVMTFVN 1301
 QY 438 -----QFLELEKTKQLRLDQV-YGNIAI-----YNFENGRRVVDGSGNWSVL 481
 DB 1302 VKDILNSRFKRENFQNV-LESDLIPYKDLTSSNYVVKDPYKFLNKEKDKFLSSYNIK 1360
 QY 482 PQQETARIIFNGKDLNLVERIAAVNPSDPL-----EYTKPDM-TLKALKIAPGF 533
 DB 1361 DSDI-----TDINFA-----NDVLGYKILSEKYSKLSLSIKYINDKQGE 1401
 QY 534 NEPN-----GNLOYQKGDITE-----FDNFQDQTSQNIKNQLAELNATNIYT 576
 DB 1402 NEKYLPLANNIETLYKTVDKIDLFVHLEAKVLNVTYEK---SNVEVKIKELN--YLKT 1456
 QY 577 VLDKIKLNARQN--ILIRKRFHYDRNN-----IAGV-ADSVVKEAHEVINSSTEGLL 628
 DB 1457 IQDKLADFKKNNFVGIADLSTDYNNHLLTKFLSTGMVFENLAKTVLSNLLDGNLQGM 1515
 QY 629 LNIQDK--IRKILSGVIVETEDTEGLKEVINDRYDMLN--SSLFQDGKTFIDPKYNDKLP 686
 DB 1516 LNIQHQCVKK-----QCQNSGCFRHLDE-----REECKLLNYKQGDKC- 1557
 QY 687 LYSNPNYKYNVAVTKENTTIINPSNG-----DTSTNGIKKI 724
 DB 1558 -VENPNPTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1590
 RESULT 4
 MSPL PLAFW STANDARD; PRT; 1639 AA.
 ID MSPL PLAFW
 AC P04933;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
 DE (PMMSA) (P195).
 OS MSP-1.
 GN Plasmodium falciparum (isolate Wellcome).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5848;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86014355; PubMed=2995820;
 RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
 RA Nicholas S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
 RA Freeman R.R.;
 RT "Primary structure of the precursor to the three major surface
 RT antigens of Plasmodium falciparum merozoites.";
 RL Nature 317:270-273(1985).
 RN [2]
 RP REVISIONS.
 RA Holder A.A.;
 RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42
 CC kDa and 19 kDa antigens which are the major surface antigens of
 CC merozoites. The maturation take place during schizont.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```
DR EMBL; X02919; CA26676.1; --
DR PIR; A24594; A24594.
DR PIR; S05603; S05603.
DR PDB; 1CEJ; 28-MAY-99.
DR Pfan; P00008; EGF; 1.
DR InterPro; IPR006209; EGF_like.
DR Malaria; Merzoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor; 3D-structure.
KW SIGNAL
FT SIGNAL 1 19
FT CHAIN 20 1639
FT CARBOHYD 116 116
FT CARBOHYD 268 268
FT CARBOHYD 764 764
FT CARBOHYD 768 768
FT CARBOHYD 783 783
FT CARBOHYD 844 844
FT CARBOHYD 920 920
FT CARBOHYD 964 964
FT CARBOHYD 1058 1058
FT CARBOHYD 1165 1165
FT CARBOHYD 1174 1174
FT CARBOHYD 1445 1445
FT CARBOHYD 1526 1526
SQ SEQUENCE 1639 AA; 187618 MW; 2C25SB6616C87F6E CRC64;

Query Match
Best Local Similarity 4.9%; Score 185.5; DB 1; Length 1639;
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

QY 5 ENRL-LNESESSQGLLYFSDLNQFAPMVVTSSTGDLSPSELENIPESENQYFOSA 63
DB 927 ENILSGKNKNIYQELIGOKSSE-NF-----YEKILDSPTFYNE 965
QY 64 IWSGFTKVKSDSYTATGADNHVTWVDDQEVINKASNS-----NKILEKRLV- 114
DB 966 SFTNFVKSADD-----INSLNDSKRRKLEIDINKLTKTLQSLFDLYNKYKLERLPD 1020
QY 115 -----QIKIQYQRENPTEKGLDFKLYWTDSONKEVIVSSDNLQPELQKQSSNRKRS 168
DB 1021 KKKTGVKYNQIKKLLKLEQESKL--NSLNPKHVL--QNFSPFFNKKKEAEATEN 1076
QY 169 TSAGPTVPDRDNGIPDSLEVEGYTVGVNKKRTFFLSPWISNHEKKGLTKY---KSSP-- 223
DB 1077 T-----LENTKILLKHY-----KGLVKYNGESSPLK 1103
QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPAR-----HPLVAAYPIVHD 265
DB 1104 TLSEESIQTEDNVASLENPKVLKLEGLKDLNLEKKKLSYLSGLHLLIA-----E 1156
QY 266 MENILSKNEDQSTQNTDSETRTISKNTSTSRHTSEVHGNAEVHASFDDIGSVYAGFS 325
DB 1157 LKEVIKKN---YTGNSPSENNT-----DVNNALESYKKFLPEGTDVATVVS 1200
QY 326 NSNSSTVAIDHSLSLAGERWAEETMGLTADTARLNANIRYVNTGTAPYINVLPPT--- 381
DB 1201 ESGSDTLEQSQPKKPASTHVGAES---NNTTSQ-NVDDEVDVLIIVPFGSESEYDDL 1256
QY 382 -SLVLGKNTLATIKADENOLSOILAPNNYPKSNLAPIA---LNQKQPSPTPTMNVN 437
DB 1257 GOVVTGEAVTFSVI---DNILSKI---ENEYEVLYLKPLAGVYRSILKQLENNVMVTFVN 1310
QY 438 -----QFLELEKTKQLRLDTQV-YGNIAI-----YNFENGVRVDTGNSWSEVL 481
DB 1311 VKDILNRFKRNFKNV-LESLLIPYKDLTSSNYVVKDPYKFLNKEKDKFLSSNYIK 1369
QY 482 PQIQETTARIIFNGKDANLVERRIAAVNPSDPL-----ETTKPDM-TLKEALKATAFQ 533
DB 1370 DSID-----TDINFA-----NDVLGYVKILSEKYKSDLSIKKYINDKQGE 1410
QY 534 NEPN---GNLOYGKDITE-----PDFNFDOQTQNTKNQALNATNIYV 576
DB 1411 NEKYLPLNLTETLYKTVNDKIDLFLVHLEAKVLNITYEK---SNVEVKIELN---YLKT 1465
```

```
QY 577 VLDKIKLNKMN--ILIRDKRPHYDRNN-----IAVG-ADESVMKEAHEVINSSTEGLL 628
DB 1466 IQDKLADFKNKNFVGLADLSTDYNNLLTFLSTGMVFNELAKTVLNLNLDGNLQGM 1524
QY 629 LNIIDKD--IRKILSGYVIBIEBTEGLKEVINRDYMLNLISSLRQDKTIDFKKNDKLP 686
DB 1525 LNIISQHCVKK-----QCPNSGCGFRHLDE-----RECKCLLNTYKQGDGKC- 1566
QY 687 LYISNPNKVNVAVTNKENTINPSENG-----DTSTNGIKKI 724
DB 1567 --VENPNPTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1599
RESULT 5
Y67 YEAST
ID Y67 YEAST STANDARD; PRT; 1658 AA.
AC Q03661; Q04988;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 187.1 kDa protein in GUA1-ERG8 intergenic region.
GN YMR219W OR YMR261.13 OR YMR959.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Church C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye K., Moule S., Odell C., Pearson D., Rajadream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII".
RL Nature 387:90-93(1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL; Z49809; CAAB9934.1; -.
DR EMBL; Z49939; CAAB9190.1; -.
DR PIR; S55101; S55101.
DR GenOnline; 142894; -.
DR SGD; S0004832; ESC1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
KW Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;

Query Match
Best Local Similarity 4.6%; Score 174; DB 1; Length 1658;
Matches 157; Conservative 140; Mismatches 278; Indels 280; Gaps 40;

QY 1 EVKQENRLNSES-----SSQGLLYFSDLNQFAPMVVTSSTGDLSPSELENIPESEN 57
DB 290 EMELEDDIDVESDAEKDSQAGTEHS-VDFSYNQPRDNTKIPVIEKYESDEHKVHQ 348
QY 58 QYFQSAWS-CFIVKKSDE-----YTPATSDN---HVTMVDVQEVINKASNSNK 105
DB 349 RYSEGDGDFGSGVNIIVDDESEDESEDSQAESANAEVYHNEHLDLDELIESDS 408
QY 106 IRLEKGRLYQIKIQYQRENPTKGLDFKL-----YWTDSQNKKEV 145
DB 409 -----ESQSAQSESGSDDDFEVKMKNEKSTSETENTSESRRQGFADATYKVK 459
QY 146 ISSDNLQIPEI-----KQSSNSRKSTSGPTVPDRDNGIPDSLEVE 190
```


DR SGD; S0003610; SMC3.
DR GO; GO:0007151; P:sporulation (sensu Saccharomyces); IMP.
DR GO; GO:0007130; P:synaptonemal complex formation; IMP.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR003405; SMC_C.
DR InterPro; IPR003395; SMC_N.
DR Pfam; PF02483; SMC_C; 1.
DR Pfam; PF02463; SMC_N; 1.
KW Mitosis; Cell cycle; Chromosome partition; ATP-binding; Coiled coil;
KW Nuclear protein.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT DOMAIN 172 482 COILED COIL (POTENTIAL).
FT DOMAIN 453 684 FLEXIBLE HINGE.
FT DOMAIN 655 1041 COILED COIL (POTENTIAL).
FT DOMAIN 1126 1161 ALA/ASP-RICH (DA-BOX).
SQ SEQUENCE 1230 AA; 141336 MW; 5152D88F7780341F CRC64;

Query Match 4.5%; Score 169.5; DB 1; Length 1230;
Best Local Similarity 19.8%; Pred. No. 0.51;
Matches 182; Conservative 136; Mismatches 355; Indels 247; Gaps 38;

QY 1 EVKQENRLNSESSESSQGLGYFDLNFQPMVVT--SSTGDLSPSSSELENIPENQ 58
DB 209 EMEQKELEKYNELERNKIYQFTLYDLRELNEVINQMERLDGYN-----NTVYSSEQ 262
QY 59 YFQ-----SAIWSGFIVKXSDYFATSDNHNHVTMVVDQEVINK--- 99
DB 263 VIQELDKREDMDIQVSKLSSI-EASLIKATDLOQAKLRESEISQKLTNNVKKIDVQ 321
QY 100 ---ASNSNKRILEKGLQYK-IOYQRENPTKGLDFKLYWTDSONKEVSISSNQLQPE 155
DB 322 QQIESNEEQNLDSDATLKEIKSIIEQRKQKSLKILPRYOELT-----KEEAMV--KLQLAS 375
QY 156 LAKSSNRKKESTGAGTPVPRDNDGIPDSLEVEGYTVDKNKTKTFLSPWISNIHE-KK 214
DB 376 LOOKORDLILKGEYARFKSKDERDTWIHSEI-----ELKSS-----IQNLNELES 422
QY 215 GUTKYSSPEKSTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAPIVHVDMENIILSKN 274
DB 423 QLOMDRTSLRKQYSAID-----EETELIDSLNGPDTKGQL-----EDFDELHLKQX 471
QY 275 EQQSTQNTDSERTTISKNTSTRTTSEVGHNAEVHASFDDIGGSVAGFNSNSSTVAI 334
DB 472 LSESLDTRKELWKEQKLTQVLETLISDVNQNR-----NVNMTMSRLANGINVKBI 525
QY 335 DHSLSLAGE-----RTWAETWG-----LNTADTARLANIRYVNTG 370
DB 526 TEKLIAPSPVFGTIGELIKVNDKYKTCAEVIGGNSLPHIVVDTEETATLNNELYRMKG 585
QY 371 TAPIVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKVLAPIALNAQKFSST 430
DB 586 GRVTF--IFPLNRLSDSDVKFSPNTTQIQTPLIKIKYEPFEKA-----VKHVFQKT 638
QY 431 PITMNTYQFLEKTKQLR---LDTDOV-YGNIAF--YNFENGRVVDVTGNWSVLPQI 484
DB 639 IVVKDLGGGLKAKKHLNAITLDGRADKRGVLTGGVLDQHKTRLESKLNLESRSQH 698
QY 485 QETTARIIFNGKDLNLVERRIAVNPS-----DPLETTKPD-MPLK 524
DB 699 KKILELDFVRNELNDITKIDQVNGNIRKVSNDRESVLTNIEVYRTSLNTKNEKLTLE 758
QY 525 EALK-IAFGNPNGLQYQKDIETEDFN----EQQTSQNIKNQLA-----E 568
DB 759 ESLNAILKLEKLNTRTFAQKLNFTFENDLLOEFDSELSKEERLESSTKE-SAAHNK 818
QY 569 LNATN-----IYTVLDKIKLNAPMILIRDKRFHYDRNNIAGV-----A 607
DB 819 LNTSDALGITTID--SLNAELESKLIPQENDLESKMSEVGDAFIFGLQDELKELQLE 876
QY 608 DRSVYKEAHREVINSST-----EGLL-----NIDKIRKI 638
DB 877 KESVEKQENAVLELGTVQVREIESIAETNNKLEKANNQORLLKLDNFQKSVKVT 936

QY 639 LSGYIIVEIEDTEGLKE-----VINDRYDM-----LN-----ISSLRQDGK 673
DB 937 MIKXTTLVTRREBELQRIREIGLLPEDALVNDPFSDDITSQQLRLNDMATEISGLKNVVK 996
QY 674 -TFIDFKKYNK-----LPLYSNFKYKNVYAVTVKENTIIINPSENG 714
DB 997 RAFENFKFKNERRKDLAERASELDESKDSIQDLVLVKLQKKQKNVAV-----DSTFQKVSFNP 1052
QY 715 DT-----STNGIKKILIFSK 729
DB 1053 EAVFERLVPRGTAKLIHRK 1072

RESULT 7
SI55 YEAST
ID -SI55 YEAST STANDARD; PRT; 1000 AA.
AC P43612;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SI54-associated protein SAPI55.
GN SAPI55 OR YFR040W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96220458; PubMed=8649382;
RA Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA Arndt K.T.;
RT "The SAP, a new family of proteins, associate and function positively
RL Mol. Cell. Biol. 16:2744-2755(1996).
RN [2]
RP SEQUENCE OF 98-1000 FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yanazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RL Saccharomyces cerevisiae";
RN Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 98-1000 FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=96287652; PubMed=8686379;
RA Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,
RA Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of
RL chromosome VI from Saccharomyces cerevisiae";
RN Yeast 12:149-167(1996).
CC -I- FUNCTION: Associates with the Slt4 phosphatase in a cell cycle
CC dependent manner. May be directly or indirectly involved in
CC Slt4-dependent functions in budding and in normal G1 cyclin
CC expression.
CC -I- PFM: Hyperphosphorylated in the absence of Slt4.
CC -I- SIMILARITY: Belongs to the SAPS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; U50560; RAC49303.1; -;
DB EMBL; D50617; BAA09279.1; -;
DB GenOnline; 140194; -;
DB SGD; S0001936; SAPI55.
DB InterPro; IPR007587; SAPS.

DR Pfam: PF04499; SAPs; 1.
KW Phosphorylation; Cell cycle.

FT CONFLICT 663
FT CONFLICT 674 689
FT CONFLICT 814 815
FT CONFLICT 818 822
SQ SEQUENCE 1000 AA; 114880 MW; AD4BF03C7C598D5 CRC64;

Query Match 4.4%; Score 166; DB 1; Length 1000;

Best Local Similarity 19.5%; Pred. No. 0.57;

Matches 173; Conservative 128; Mismatches 323; Indels 262; Gaps 42;

QY 1 EVKQENLLNESQGLGYFS-DL-----NFOAPVMTSTTGD-----LSIPS 47
DB 90 EKEEERSNSESSTTSPSSGSTKTDLDEEDISNATAPMVYTNLDNFIERMLVET 149
QY 48 SELENIPSNQYFQAGNSGFIKVKSDSYTFATSDADNHTVWVDD-GEVINKASNSKI 106
DB 150 ELLNELSRQKTLDFICGFFDKTKNKV-----NNMEYLVQLMCEISKIKTATV 203
QY 107 RLEKGLYQIKYQRENTEKGLDFKLYWTDSQNKKEVSSDNLQLPKXKSNRKK 166
DB 204 DLNN-----LDYQEQQLD-----DSSQEDVYVESDTQEKEEDNNNNKK 247
QY 167 R-----STGAGTVPDRDNDGIPDSLEVBGYTVVKNKRTFLSP-----WISNIHERKG 215
DB 248 RKRGGSSFGNDINNDDDDANEDDESAYL---TKATIIEIFSLDWLIS-----ES 299
QY 216 LKYKSSPK-KSTASDPYSDPKYTKGRIDKXVSEARHPLVAAPYIVHVDMENTILSKN 274
DB 300 LKVNQSYLNKIWSINQP-----NFNSE-NSPLV---PIFLKINQNLTLR- 341
QY 275 EDQ-----STQWTDSET-----RTISKNT 293
DB 342 QOYLNFIFERSFVDDMLKHVDISLMDFFLKILSTOKIESPTGIIELVYDQNLISKL 401
QY 294 STSRTHSEVHGAENVHAFPIGGSVAGFNSNSTVAL-DHSL--LAGERTWATM 350
DB 402 SFLNNKESP-----ADIQACVGDPLKALIAISANAPLDDISGPNLSLTLQASPEIAKL 457
QY 351 GLNTADTARNANIRVYVGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNY 410
DB 458 DI-----MINQGAALNT-----TVSIVI---ELIRKNSDDVQNLTLTTIKT 498
QY 411 YPSKNLAPIALNAQKDFSTPTIMYNOFLEKTKQLRLDQDVYGNIAITYNFENGRV 470
DB 499 HPPSNRDPYLVGLLRKFNHLSDFPQIILDIENDANIPHENQLHERFKPLGFERFKV- 557
QY 471 VDTGSNWSVLPQIQETTARIIFNGKDLNVERRIAANVP--SDPLETTKPDWTL--KEA 526
DB 558 -----VELIAELLHCSNGLMNSKRAERIAARRDKVRSQSLSHLDALNDSIEKEQ 610
QY 527 LKIAFG-----FNEPNGNLQYQKDI-----TEPDPNFD---OQTSQNIKNQLAE 568
DB 611 LKTKHSPTRTDHLKNNNGKIDNNDNDDSDYDEIDSEFEPYINMKQNIKLRTDP 670
QY 569 LNATNIVYLDKILNAKNMILRDKRFYDRNNAIAGAD-----ESVYKEAHEVIN-- 621
DB 671 TVGT-----YKSNMYIRD---CFQNNELFTHPWNFWNFWNFIIDIIQIFNGR 716
QY 622 ---SSTEGLLNID--KDIRKILSGYVIEE---DTEGLKEVINDRYDMLNLS---LR- 669
DB 717 MDSYNSFLVLISLNLKSSYQFMTDIVISEKGTDVRSRFPVDRDNEDFKITTDIFLRG 776
QY 670 QGKTFIDPKKN-----DKLPLYISN-----PNYKNVY 699
DB 777 YQDSYKFEYLRKNNLGMVGHVILAEVWFKSLYKVDYISRYSSNLQTEWQYVSEEV 836
QY 700 AVTK-----ENTIIINPSENGDSTN 719
DB 837 NEIRMYSKILGGGSYIDGNGNLIIPQLPDNTVILTP--NGDASN 880

RESULT 8

WAPA_BACSU STANDARD; PRT; 2334 AA.
ID WAPA_BACSU
AC Q07833;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Wall-associated protein precursor.
GN WAPA OR N17G OR BS039230.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93302506; PubMed=8316082;
RA Foster S.J.;
RT "Molecular analysis of three major wall-associated proteins of
RT Bacillus subtilis 168: evidence for processing of the product of a
RT gene encoding a 258 kDa precursor two-domain ligand-binding
RT protein.";
RL Mol. Microbiol. 8:299-310(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=95219088; PubMed=7704263;
RA Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;
RT "Cloning and sequencing of a 29 kb region of the Bacillus subtilis
RT genome containing the hut and wapa loci.";
RL Microbiology 141:337-343(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=97124196; PubMed=8969509;
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT containing the lic and cel loci, and creation of a 177 kb contig
RT covering the gnt-sacXY region.";
RL Microbiology 142:3113-3123(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Enrian K.D., Errington J., Fabbri C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Rullo M.P., Itaya M., Jones L.,
RA Joris B., Kazamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Melillo R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takanashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vanderbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengruber T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

"The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.";
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,
 CC MOTILITY, SECRETION OR DIFFERENTIATION.
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED
 CC INTO THE MEDIUM.
 CC -!- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE
 CC 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE
 CC C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED
 CC MOTIF REPEATED 31 TIMES.
 CC -!- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME
 CC SIMILARITY TO THE REPEAT IN *E. COLI* RHS GROUP OF PROTEINS (RHS-A-D).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; L05634; AAA22883.1; -
 CC EMBL; D31856; BAA06656.1; -
 CC EMBL; D29985; BAA06260.1; -
 CC EMBL; D83026; BAA11683.1; -
 CC EMBL; Z99124; CAB15959.1; -
 CC PIR; S32920; S32920.
 CC Subtilisin; BG10797; wapa.
 CC InterPro; IPR003305; CBM_CenC.
 CC InterPro; IPR006530; YD.
 CC Pfam; PF02018; CBM 4_9; 1.
 CC Pfam; PF05593; RHS repeat; 14.
 CC TIGRfams; TIGR01643; YD repeat; 2x; 17.
 CC Call wall; Repeat; Signal; Complete proteome.
 CC SIGNAL
 CC 1 28
 CC OR 32 (POTENTIAL)
 CC WALL-ASSOCIATED PROTEIN.
 CC 3 X 101 AA APPROXIMATE TANDEM REPEATS.
 CC
 CC 1 1
 CC 2 2
 CC 3 3
 CC 4 4
 CC 5 5
 CC 6 6
 CC 7 7
 CC 8 8
 CC 9 9
 CC 10 10
 CC 11 11
 CC 12 12
 CC 13 13
 CC 14 14
 CC 15 15
 CC 16 16
 CC 17 17
 CC 18 18
 CC 19 19
 CC 20 20
 CC 21 21
 CC 22 22
 CC 23 23
 CC 24 24 (APPROXIMATE).
 CC 25 25
 CC 26 26
 CC 27 27
 CC 28 28
 CC 29 29
 CC 30 30

FT REPEAT 2120 2139 2-31.
 SQ SEQUENCE 2334 AA; 258329 MW; B75138CCD278BAA3 CRC64;
 Query Match 4.4%; Score 165.5; DB 1; Length 2334;
 Best Local Similarity 18.2%; Pred. No. 1.9;
 Matches 173; Conservative 126; Mismatches 336; Indels 317; Gaps 42;
 Qy 1 EVKQERLLNESSSSQGLLYGYPFDLNFQA-----PMVTSSTTGDLSIPSELENIP 54
 Db 269 EVERSCKVSKLEKNEBEGYLLHLTADEMLKDPERYVPSIDPSTLSVSSDTFVMSAYP 328
 Qy 55 SENQYFQSAIWSGFIK--VKKSDEYTFATSDAHVHTM----- 89
 Db 329 TTVYASSQKWDANLKAAYLVKGTGYD-KTGTNYAFMKFNLLKPIQNMVTYKATLKTYVA 387
 Qy 90 -----WDDDEVINK-----ASNSKIELEGR--LYQIKIQYOR 122
 Db 388 HSYVGTGKATGLWD--TVNSYDNDNAKVTWNTKPSKNGIKADVHGQWASVDVTAARKS 444
 Qy 123 ENPTEKGLDFKLYWTDSONK---KEVISSDNLQPELKQKSSNRKRSSTAGTVP--- 176
 Db 445 WNSGANYGFKLH-TNGNGKEYWKLIJ-----SANSANKPYIEVTYTPKGN 491
 Qy 177 -----DRDNDGIPP-----SLE-VEGYTVVQNKRTFLSPWISNIHEKKGLTKYKSSP 223
 Db 492 TPTIKAYHNGDSTGYFDISWKEGAKGYKWTYNGKEYQALISAGNV-----TSMWTKG 545
 Qy 224 EK-WSTASDPYSDPEKVTGRID-KNVSPEARHPLVAAYPIVHVDMENIILSKNE----- 275
 Db 546 KKIWPTSAEIAASKRYKL--HLDGDKGAELALDP-----SPVYKNSGGSYATSKNYWGVSA 599
 Qy 276 --DQSTONTSETRTTSKNTSTRTWTSVHGNAEVAHAFDFIGGSVAG-----FS 325
 Db 600 IFDQEGAMGAPAKVPVPGVKAQAPSKYNNNGNA-TGYFDLSKAVSGATGYKQVQVFN 658
 Qy 326 NSNSSTVAIDH--SLSLAGERTWAETMGLNTADTA-----RLNANIRVYNTGTA 372
 Db 659 KGFEITLDLGNQTSWTKGKIWIPTSABIKAGKVALHLKDGSAEPIFNGPYKAGGD 718
 Qy 373 PIYNVLTSLVLGRKNTLATIKADENQLSQILAPNN-----YPSK-----NLA--- 417
 Db 719 GAKNYSFKIIAYNKDGEALASPAATPALDIAKPNKVTGYLTNTKSSQTGYVNLWEK 778
 Qy 418 -----PIALNAQKDESSPTI-----TMVYNQFLELEKTK-----CLRL 450
 Db 779 VQNAKGYKVIYNGKEYQSFVGDADHWTTONKNIMPTSEIKAGSYKLTGDKGBELAL 838
 Qy 451 DTDQVYGN-----IATYN--FENGVRVVDTSNWSSEV-L 481
 Db 839 DSPVYNNANGYKKNYSFTLVAYDANGETIPTAFPFTFHEGAFLEGTEEYWSIID 898
 Qy 482 P--QIQTETARIIFNGKDLNJ-----VERRIAAVNPSDPL-----ETT----- 517
 Db 899 PSGQLNGATGNVIYNEEDLSIDGRGPGGLSRTYNSLDSDHLPFGQGWYADAETSVISTD 958
 Qy 518 -----KPDMTLKEALKIAPG-FNEPNGNLQYQOKDITEFDNFDDQTSQNIKQLAEL- 569
 Db 959 GGAMVIDEDATTHRTKADGTQYPTGVYLELTETADQFILTKDQNTAYFNKKGKGLQ 1018
 Qy 570 -----NATNIYTVLDIKLNKAKNIIIRDKRHFYDRN-----NIAVGADESV 611
 Db 1019 KVDGHNNAVTYTYNDKNQLTAITDASGRKLTFTYDENGHVTSITGPKNKVTVSYENDL 1078
 Qy 612 VKE-----AHR--EVINSSTEGILL 629
 Db 1079 LKKVTDGTVTSYDYDSEGLVQYQSANSTEAKPVEFTEYQYSHRLEKAINAKETVYV 1138
 Qy 630 NIDKDI-----EKILSGY-----IVEIEDTEGLKEVINDRVMDLNI 665
 Db 1139 SYDADKKTLLMTQPNKRVQYGYNEAGNPQIWDABGLKITNTKYEKNV 1190

RESULT 9


```

RX MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT truncated head of Dictyostelium discoideum myosin to 2.7-A
RT resolution.";
RL Biochemistry 34:8973-8981(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the
RT Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
RL Biochemistry 35:5404-5417(1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RX MEDLINE=97452860; PubMed=9305951;
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT "X-ray structures of the MgADP, MgATPgammaS, and MgAMP-PNP complexes
RT of the Dictyostelium discoideum myosin motor domain.";
RL Biochemistry 36:11619-11628(1997).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RX MEDLINE=98070605; PubMed=9405148;
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT "X-ray crystal structure and solution fluorescence characterization
RT of Mg2(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RT Dictyostelium discoideum myosin motor domain.";
RL J. Mol. Biol. 274:394-407(1997).
CC -!- FUNCTION: Myosin is a protein that binds to actin and has ATPase
CC activity that is activated by actin.
CC -!- SUBUNIT: Myosin II heavy chain is two-headed. It self-assembles
CC into filaments. Hexamer of 2 heavy chain subunits (MHC), 2 alkali
CC light chain subunits (MLC) and 2 regulatory light chain subunits
CC (MLC-2).
CC -!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC CORTX.
CC -!- DOMAIN: Each myosin heavy chain can be split into 1 light
CC meromyosin (LMM) and 1 heavy meromyosin (HMM). It can be further
CC split into 2 globular subfragments (S1) and 1 rod-shaped
CC subfragment (S2).
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC characteristic for alpha-helical coiled coils.
CC -!- PTM: Phosphorylation inhibits thick filament formation and reduces
CC the actin-activated ATPase activity.
CC -!- MISCELLANEOUS: Dictyostelium myosin II has no K(2)EDTA ATPase
CC activity, perhaps correlated with the absence of a Cys at the SH-1
CC position (688).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14628; AAA33227.1; -.
CC PIR; A26655; A26655.
CC PDB; 1MMA; 03-DEC-97.
CC PDB; 1MMD; 17-AUG-96.
CC PDB; 1MNG; 03-DEC-97.
CC PDB; 1MMN; 03-DEC-97.
CC PDB; 1MND; 17-AUG-96.
CC PDB; 1MNE; 17-AUG-96.
CC PDB; 1VOM; 23-DEC-96.
CC PDB; 1LVK; 28-JAN-98.
CC PDB; 1DOX; 20-DEC-00.
CC PDB; 1DOY; 20-DEC-00.
CC PDB; 1DOZ; 20-DEC-00.
CC PDB; 1D1A; 20-DEC-00.

```

```

DR PDB; 1D1B; 20-DEC-00.
DR PDB; 1D1C; 20-DEC-00.
DR PDB; 1FMV; 20-DEC-00.
DR PDB; 1FMW; 20-DEC-00.
DR PDB; 1G8X; 17-JAN-01.
DR PDB; 1JMY; 07-NOV-01.
DR PDB; 1JX2; 07-NOV-01.
DR DictyBase; DDB0002015; mbca.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR008989; Myosin_S1_N.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00663; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSc; 1.
DR PROSITE; PS50096; IQ; 1.
DR Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
KW Calmodulin-binding; Methylation; Phosphorylation.
KW MYOSIN HEAD-LIKE.
FT DOMAIN 1 761
FT DOMAIN 762 791
FT DOMAIN 817 2116
FT NP_BIND 179 186
FT DOMAIN 638 660
FT DOMAIN 738 752
FT MOD_RES 130 130
FT MOD_RES 1823 1823
FT MOD_RES 1833 1833
FT MOD_RES 2029 2029
FT TURN 3 5
FT TURN 7 8
FT HELIX 10 15
FT HELIX 25 28
FT STRAND 34 37
FT TURN 41 42
FT STRAND 48 55
FT STRAND 59 63
FT STRAND 69 73
FT HELIX 74 76
FT STRAND 78 79
FT HELIX 83 85
FT TURN 86 87
FT STRAND 90 90
FT HELIX 91 93
FT HELIX 99 110
FT TURN 111 113
FT STRAND 116 119
FT TURN 120 121
FT STRAND 122 126
FT HELIX 137 142
FT TURN 143 145
FT HELIX 148 150
FT HELIX 155 169
FT TURN 170 170
FT STRAND 173 179
FT TURN 181 182
FT HELIX 185 200
FT HELIX 210 226
FT STRAND 227 228
FT TURN 234 235
FT STRAND 236 237
FT STRAND 240 247
FT TURN 249 250
FT STRAND 253 261
FT HELIX 265 268
FT TURN 269 269
FT TURN 273 274
FT STRAND 278 278
FT HELIX 279 287
FT HELIX 290 296

```

FT TURN 297 297
 FT HELIX 301 303
 FT TURN 305 307
 FT TURN 316 317
 FT HELIX 320 334
 FT TURN 335 335
 FT HELIX 338 355
 FT TURN 356 356
 FT STRAND 360 360
 FT STRAND 368 368
 FT HELIX 373 382

Query Match 4.3%; Score 162.5; DB 1; Length 2116;
 Best Local Similarity 19.4%; Pred. No. 2.3;
 Matches 138; Conservative 113; Mismatches 219; Indels 243; Gaps 33;

QY 86 HVTWVDDOEVINKASNSKIRLEKGLYQIKYQORENPTKGLDFKLYWTDSONKKE- 144
 DB HVNQLSEEEKQKESNEKRVKVDLEK-EVSELKQIEEVASKAV-----TEANKKES 1346
 QY 145 -----VSSDMLQPELKQKSSNRKSTSGPTVPRNDGIPDSLEVEGYT 193
 DB ELDEIKRQYADVWSRDKSVLEQLTKLOAKNEELRNTA-----BEAQGLDRAERS--- 1396
 QY 194 VDVGNKRTF-LSPWISNHE-----KKGLTKYKSSPEKWSSTADPVSPFKEYT 240
 DB ---KKAEFFDEEAVKLEBETAKVKYKAKMKAETDYRSTKSELDDAKNVSSQYVQI 1453
 QY 241 GRIDKNVS-----PEARHPLVAAPYVHVDMNIIILSKNEDOSTONTSETRTTSK 291
 DB KRLNEELSELRSVLSEADERCNSAKAKKTAESALESL---KDEIDAAANAKAKERSSK 1510
 QY 292 NTS*SRHTSHSVHNAEVAHAFPDIGSVAGSFGNSNSTVAIDHLSLAGERTWAETWG 351
 DB ---EVR-VASELESLEDKSGTVNVEIRKDD-----AEIDDD 1544
 QY 352 INTADTARLNANIRVYNTGAPIYVLPSTSLVGLKGNQIATIKADENGLSQILAPNNY 411
 DB LR-----ARLD-----RETESRIKSDERK----- 1563
 QY 412 PSKNLAFIALNAQDFS--STPITMNYNQFLEKTKQLRLDQVYGNIAVYNFENGRV 469
 DB ---KNTRQKQFADLEAKVEEAQREVTTIDRLKX-KLESDDI--DLST----- 1603
 QY 470 RVDTCNSWSEVLPQIQTARI-IFNCK---DNLVBERIAAVNPDPLETKPDWTLKE 525
 DB QLDT-----ETKGRIKTEKSKKLEQTLAERRAAEEGSSKAAD-----EE 1643
 QY 526 ALKIAFGFNEPNGNLQYQGDITDFDNFDQ-----TSQNIKNQALNATNIYTVLD 579
 DB IRKQVW-----QEVDELRAQLDSERAAALNASEKKIKSLVAE-----VD 1681
 QY 580 KIKLNKANNILIRDKRHYDRN---NTAVGADESIVKEAHREVINSSTGLLLNDKQJR 636
 DB EVKQLEDEILAKOKLVAKRAULEVELEVRDQLEEDSRSELEDSKRLTTEVE-DIK 1740
 QY 637 KILSGYIVIEE-----DPEGLKEVINDRYDMINTSS-----LRQDGKTFI- 676
 DB 1741 K---KYDAEVEQNTKLEAKKLTDDVDTLKQLEDEKKKLNSEKAKKRLSENEDEFLA 1797
 QY 677 -----DFKYNKDLPLYISNPNYKVNVTAVTKNTIINPSENGD 715
 DB 1798 KLDIAEVKNRSRAEKDRKKYKNDL-----KDTYKLNDEAATKTQTEIGAALKE 1846

RESULT 11

SPOF SCHPO
 ID SPOF SCHPO STANDARD; PRT: 1957 AA.
 AC Q10411.Q3US9;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Sporulation-specific protein 15.

EMBL: Z70690; CAA94624.1; -
 EMBL: AB027811; BAB87115.1; -
 PIR: T38077; T38077.
 GeneB Spombe; SPAC13.06c; -
 Sporulation; Coiled coil.
 DOMAIN 199 785 COILED COIL (POTENTIAL).
 DOMAIN 804 1235 COILED COIL (POTENTIAL).

SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
 MEDLINE=20107136; PubMed=10639340;
 Ikemoto S., Nakamura T., Kubo M., Shimoda C.;
 "S. pombe sporulation-specific coiled-coil protein Spo15p is localized
 to the spindle pole body and essential for its modification.";
 J. Cell Sci. 113:545-554 (2000).
 [2]
 SEQUENCE FROM N.A.
 STRAIN=972;
 MEDLINE=21848401; PubMed=11859360;
 Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 Rutherford K., Ruster S., Saunders D., Seeger K., Sharp S.,
 Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 Meltjens I., Vanterre E., Rieger M., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 Geffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
 Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 Sipakowski G.V., Ussery D., Barrell B.G., Nurse P.;
 "The genome sequence of Schizosaccharomyces pombe.";
 Nature 415:871-880 (2002).
 [3]
 SEQUENCE OF 705-871 FROM N.A.
 STRAIN=968 E90;
 MEDLINE=20233868; PubMed=10759889;
 Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
 Hiraoka Y.;
 "Large-scale screening of intracellular protein localization in living
 fission yeast cells by the use of a GFP-fusion genomic DNA library.";
 Genes Cells 5:169-190 (2000).
 CC -!- FUNCTION: Has a role in the initiation of spore membrane
 formation.
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Spindle pole body.
 CC -!- SIMILARITY: Belongs to the MPC70 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z70690; CAA94624.1; -
 EMBL: AB027811; BAB87115.1; -
 PIR: T38077; T38077.
 GeneB Spombe; SPAC13.06c; -
 Sporulation; Coiled coil.
 DOMAIN 199 785 COILED COIL (POTENTIAL).
 DOMAIN 804 1235 COILED COIL (POTENTIAL).

FT DOMAIN 1320 1471 COILED COIL (POTENTIAL).
 FT DOMAIN 1481 1723 COILED COIL (POTENTIAL).
 SQ SEQUENCE 1957 AA; 22785 MW; 3F480CA0617D9DA CRC64;
 Query Match
 Best Local Similarity 20.4%; Score 162; DB 1; Length 1957;
 Matches 177; Conservative 132; Mismatches 328; Indels 232; Gaps 40;
 QY 2 VQENRLLNESE-----SSOGLLYGFSDLNFO---APMVTSSTTGDLSIPSE 49
 Db 129 VTOKSNLLNELQVRSKLAALSHENGILSLQSSNKKDKMTSSVTTLTSEEDVSYPQK 188
 QY 50 LENIPSENYFOSAIWS---GRIKY-----KASDEYTFATADNHVTVWDDQVINKAS 101
 Db 189 LTMESNFGAKSEAYDLSRQLLTTEKLDKKEKDYKIKEDVSISIKASLAEEQASNKL 248
 QY 102 NSNKIRLEK---GRLYQIKIQRENPTPE---KGLDFPLYWTDSONKEVISSDNLQPE 125
 Db 249 RGEQERLEKLVSSNKTVSTLQNTENSLEAECKTLQEKL-----EKCANEEDSKLEE 302
 QY 156 LKQKSN-----SRKSTSGAGTVDPDRDNGDIPDSLEVEGYTVVKNKR----- 200
 Db 303 LKHNVANYGDAIYHDKLTJEDLSRISBFDN-----LKSRDTLSIKNEKLEKLENTI 356
 QY 201 -TFLSPWISNIHEKKGKLTAKYS-----PEKWTASDPYSDEK-----VTGRID----- 244
 Db 357 GSKDSRTSNQLEEEWVELKESNTHISQLTDAESKLSPEQENKSLKSIDYQNNLS 416
 QY 245 -----KQVS-----PBARPLVAA-----YPIVHVM----- 266
 Db 417 SKDKMVKQVSSOLEEASLAHATKLABINSERDFONKKIKDFEIQDLRACLNSEN 476
 QY 267 ---ENILSKNEQDSQNTD---SSTRISKNTSTS-----RTHSEVHGNAEVHAFPD 315
 Db 477 ELKEKALIDKQDELANNLEQIKQKQVSESTQSSLOSLOLDILNEKKKEHVESQNE 536
 QY 316 IGSVSAGSNGNS-----STVAIDHSLSLAGERTWAGTMTGLNTADTARLNNRYNTG 370
 Db 537 LKGELOTEISNEHLSQSLTLAAEKAATNNELSESK---NSLQTL-CNA----- 585
 QY 371 TAPIYVNLPTSLVLGKQKQ-----TLATIKADENQSLQILAPNYPVPSKQLAPIALNAQD 426
 Db 586 ---FQEKAKSVMLKENQNFSSLDTSFKKLNESHQLENNHQTITKQL-----KD 634
 QY 427 FSSPTITMYNPFLEKTKQLRLTDQVYGNATYFNENGRVRVDTGSNWSVLPQIE 486
 Db 635 TSS-----KIQQLERANFEQKESLSDENNDLRK-----LLKLEE 672
 QY 487 TTARIIFNGKDLNLVERRIAAVNPSPLETTKPDN-TLKEALKI-----AFGNFPGNLQ 541
 Db 673 SNKSLIKQEDVDSLEKNI-----QTLKEDLRKSEALRFSKLEAKNLEVIDNL- 722
 QY 542 YQKQDIT-EFDNFDQOTSQNTQNLQALBNATNIYTVLDKIKLANONILIRDKRFYDR 600
 Db 723 -KGKHETLEAQRNDLHSSLSDAKNTAILSSSLTKSSEDDVKRLTANVETLTQDSK----- 776
 QY 601 NNTAVGADESVMKEAHREVINSFTEGLLNLDKQIR-----KILSGYVIE-----DT 649
 Db 777 -----AMKQSFSLVNSVQS---ISNLYHELRDDHVMQSQNTLLSESKLTDK 824
 QY 650 EBLKE---VINDRYDMNLTSSLRQCKFTPIDPKYNDKLPYISNPYKNVNYAVTKENT 706
 Db 825 ENLTQOMTLLIDNVQKLEKHVNSKSV-SELKEYNGKLSLDLKNLSSLV-AISNDQ 882
 QY 707 IINP-----SENGDT-----STNGIKKI 724
 Db 883 ILTQLAELSKQVDSLEQESAQLNSGLKSL 911

RESULT 12
 P115_MYCHR
 ID P115_MYCHR
 AC P41508;

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE P115 protein.
 OS Mycoplasma hyorhinis.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2100;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91113990; PubMed=1825306;
 RA Notariccola S.M., McIntosh M.A., Wise K.S.;
 RT "A Mycoplasma hyorhinis protein with sequence similarities to
 nucleotide-binding enzymes.";
 RL Gene 97:77-85(1991).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- DOMAIN: Consists of two putative central coiled-coil regions
 flanked by putative globular regions at the N- and C-termini.
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYOPLASMA
 P115.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M34956; AAA25423.1; -
 DR PIR; JQ0894; JQ0894.
 DR InterPro; IPR003439; ABC transporter.
 DR InterPro; IPR005289; GTP-binding_dom.
 DR InterPro; IPR003405; SMC_C.
 DR InterPro; IPR003395; SMC_N.
 DR Pfam; PF02483; SMC_C; 1.
 DR Pfam; PF02463; SMC_N; 1.
 DR TIGRFAM; TIGR00850; MG442; 2.
 KW ATP-binding; Coiled coil.
 FT NP BIND 32 39
 FT DOMAIN 169 224
 COILED COIL (POTENTIAL).
 FT DOMAIN 231 400
 COILED COIL (POTENTIAL).
 FT DOMAIN 569 821
 COILED COIL (POTENTIAL).
 FT DOMAIN 884 912
 ALA/ASP-RICH (DA-BOX).
 SQ SEQUENCE 979 AA; 110566 MW; 30D51C56B56280F4 CRC64;
 Query Match 4.3%; Score 161.5; DB 1; Length 979;
 Best Local Similarity 18.9%; Pred. No. 0.92;
 Matches 175; Conservative 132; Mismatches 342; Indels 277; Gaps 40;
 QY 13 BSSOGLLYGFSDLNFOAPMVV-----TSSTTGLDLSIPSELENIPSE 56
 Db 51 EQSAKQLRGLNMDVIFAGSKTVKQKAMVKLTFKNEIDAETKQIFTISRLKRCQGT 110
 QY 57 NOYFQS-----AIWGFIVKKSDEYTFATADNHVTVWDDQD--VINKAS 102
 Db 111 NEFYNDQPVRYKQIKNLAVESG---ISKSSLAISQGTISEIAEATPEQKAVIEBAAG 167
 QY 103 SNKILEKGRLYQIKIQRENPTPEKGLDFPLYWTDSONKEVISSDNLQPELKQSSN 162
 Db 168 TSKYKLDKEEQAQKLIR-----TNDALD-KLQAIKELEQVNSLDK-QASKAKIVLEK 219
 QY 163 SKKKESTSGAGTVDPDR-----NDGIPDSLEVEGYTVVVK-NKRTFLSPWISNIHEKKG 215
 Db 220 SKALESVEVGLIVNDLNFNEKLNNTSLLEVEQQRNDLELNTQTYESSISQTVH----- 275
 QY 216 LTKYKSSPEKWTASDPYSDEKVTGRID--KNVSP-----EARRHPLVAAYPIVHV 264
 Db 276 ---FKTEV-----SSIQEITSKLDNLKNALEINLQEARIERKRLIISGEIV-V 322
 QY 265 DMENII--LSKN-EQOSTQNTDSETRISKNTSTSRTH----- 300
 Db 323 DOKTKIEBKQVSESLKIQINASKQRETELDQQLTRLNAYANSKLQENDINKEIGVLE 382

QY 301 --SEVHGAETHASFEIDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAE---TWGLNTA 355
 Db 383 KSAANAANILKQFENKFSKGIKTIKNGFLDGYIGLSELFKEVSEFSLAETV 442
 QY 356 DPARLN-----ANIRYVNTGTAPINYVLPFTSL-----VLG 386
 Db 443 LGAALNQIVMTKSDVLOAIDFKKNSLKGATFIPLTISKEREVREDHLLVLKQKQGLG 502
 QY 387 KNOTLATIKADENQ-----LSOILAPNYYPSKNLAPI-----ALN 422
 Db 503 VAKELIEFTQNKLFGLGILNVLVDVNDANRIAKILDKHYTIVSLEGLFRPGGTIT 562
 QY 423 AOKDFSSTPITWYNYQFLEKTKQLRLDTQ-----VYGNATYVFNFGVRVD 472
 Db 563 GSKGLERTSI-LNYDIKIK-EHTNLTKEADQIHLKIKQOITVNEISTVSTIQVKIE 620
 QY 473 TGS-----NWSEVLPQIOE--TTARILFNGKDLNVERIAAV----- 508
 Db 621 ANSINKLNILNEELNNKLNASEIFKEQOEDQESLNSFPSEKLN-IEKQISTLIETLN 679
 QY 509 NPSDEL-----ETTKPDMTLK-----EALKIAGFNEPNGNLQYQK 545
 Db 680 SKKRLTNLISQKGEYKQKQELDAKLKLNTOHSDSITEQNRKFLVEQKQKLSHYK 739
 QY 546 DITE-----PDFPDQOTS--QNIKNQLABLNATNTYVLDKIKKNAMNLIIRKRF 596
 Db 740 LYLEASPOYSLDLDIEQARHFVDSLKKELGNNLEAITEFEVNR-----YQBKKQ 795
 QY 597 HYDRNNAVGADESVVYKHAHREVINSSTEGLLNIDKD-----IRKILSGYIVIEDTEG 651
 Db 796 YIEELTTAKSKEIEAISLDDKIINKTTE--IVNLVNEFNVPQKMGKGAKIHT-- 851
 QY 652 LKEVINDRYMLNIS---SLRODGTFTDFKKYND-----KLPIYI- 689
 Db 852 -----DKNDILNSGVEISAQPGXTIKNLRIFSGGERAKIIAISLLFAILKARPIPLCIL 905
 QY 690 -----SNPNYKNVYAVTKENT 706
 Db 906 DEVERALDESNIYRVEFLKIKENT 931

RESULT 13
 BDX_CLOBO
 AC BDX_CLOBO STANDARD; PRT; 1276 AA.
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)
 DE (Bontoxilysin D).
 GN BOTD.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BVD/-3;
 RX MEDLINE=91016853; PubMed=2216736;
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
 RA Kozaki S., Kriegstein K., Henschen A., Gill D.M., Niemann H.;
 RT "Nucleotide sequence of the gene encoding Clostridium botulinum
 RT neurotoxin type D.";
 RL Nucleic Acids Res. 18:5556-5556(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CB16;
 RX MEDLINE=93042276; PubMed=1420572;
 RA Sunagawa H., Ohnaya T., Watanabe T., Inoue K.;
 RA "The complete amino acid sequence of the Clostridium botulinum type D
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RT phase d-16 phi genome.";
 RL J. Vet. Med. Sci. 54:905-913(1992).

RNA PARTIAL SEQUENCE.
 RC STRAIN=D-SA, and D-1873;
 RX MEDLINE=89339741; PubMed=2668193;
 RA Morifishi K., Syuto B., Kubo S., Oguma K.;
 RT "Molecular diversity of neurotoxins from Clostridium botulinum type D
 RT strains";
 RL Infect. Immun. 57:2886-2891(1989).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94230352; PubMed=8175689;
 RA Yamaaki S., Baumbast A., Binz T., Blas J., Link E., Cornille F.,
 RA Roques B., Fyke E.M., Suedhof T.C., Jahn R., Niemann H.;
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
 RT F botulinum neurotoxins and tetanus toxin.";
 RL J. Biol. Chem. 269:12764-12772(1994).
 CC -!- FUNCTION: BOTULINUM TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-LEU-61 BOND OF
 CC SYNAPTOSOMES-1 AND -2.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (by similarity).
 CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H). The light chain has the pharmacological activity,
 CC while the N- and C-terminal of the heavy chain mediate channel
 CC formation and toxin binding, respectively.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -!- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 CC BACTERIOPHAGE.
 CC -!- SIMILARITY: Belongs to peptidase family M27.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X54254; CAA38175.1; -;
 CC EMBL; S49407; AAB24244.1; -;
 CC PIR; S11455; S11455.
 CC HSP; P10845; 3BTA.
 CC MEROPS; M27.002;
 CC InterPro; IPR008985; ConA like lec.gl.
 CC InterPro; IPR002160; Kunitz legume.
 CC InterPro; IPR006025; Pept M Zn BS.
 CC InterPro; IPR000395; Peptidase M27.
 CC Pfam; PF01742; Peptidase M27; 1.
 CC PRINTS; PR00760; BONTOTOXILYSIN.
 CC ProDom; PD001863; Bontoxilysin; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT CHAIN 1 442 BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.
 FT METAL 443 1276 BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.
 FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 230 230 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 437 450 INTERCHAIN (PROBABLE).
 FT VARIANT 15 16 ND -> PV (IN STRAIN D-1873).
 FT VARIANT 17 18 ND -> LQ (IN STRAIN D-1873).
 FT VARIANT 452 452 K -> Q (IN STRAIN D-SA).
 FT VARIANT 457 457 R -> T (IN STRAIN D-SA).
 FT VARIANT 457 457 R -> F (IN STRAIN D-1873).
 FT VARIANT 462 462 A -> D (IN STRAIN D-1873).


```
Db 2308 YRNISETKLQMEHSTDFKPMELHKGNET-----NNKSLEKEKLLKSVN--DHMH 2359
Qy 516 YTKPDWTLKALKI-----AFGFPENGNLQYQKQITEPDFNF--DOOTSQNIKNQLAEL 569
Db 2360 SMEAEM-IRKGLKYTPESVQNNINNYSVIEAEVKTLEEIDRDYGDNYQVIEEHKKQFSIL 2418
Qy 570 -NATNIYTVLDKIKLNAMKMLIRDRKPHY---DRNNIAGADESVVKEAHREVINSST- 624
Db 2419 IDRTN--ALMDDHIFKKN-----NYNLMVNTETTHRVNDVIEKITNKLVOAKTE 2469
Qy 625 -EGLLNI---DKDIRKILSGYVIEIDTEGLKE---VINDRYDM-----LNISLR 669
Db 2469 YEILENIKQNDMLQNLIFLKVYSIIEYFENVKKESILNDLYEQERLLKIGEHLEIK 2528
Qy 670 QDGKTFDFKXNDKPLYSN-----PNYKNVYATVKTENTLINS-----ENGDS 717
Db 2529 RVNTELUSSYEIOQWEMSKNLLKKSQNNY-TSIYELEREANEINDAKQKDDDTI 2587
Qy 718 TNGIKKILIFSKG 731
Db 2588 LNSVLEAAI-QKRG 2600

RESULT 15
FAT1_SCHPO STANDARD; PRT; 1385 AA.
AC OI3735; Q9UTJ0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Actin interacting protein 3 homolog.
GN FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Gaffeau A., Cadieu E., Dreano S., Gloux S., Delaure V., Mottier S.,
RA Calibert F., Aves S., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipakovski G.V., Usery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880 (2002).
RN [2]
RP IDENTIFICATION, AND GENE NAME.
RX MEDLINE=20143585; PubMed=10679021;
RA Jin H., Amberg D.C.;
RT "The secretory pathway mediates localization of the cell polarity
```

```
RT regulator Atp3p/Bud6p.";
RL Mol. Biol. Cell 11:647-661 (2000).
CC -!- FUNCTION: Involved in the organization and/or function of the
CC actin cytoskeleton.
CC -!- SIMILARITY: TO YEAST BUD6.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z97208; CAB10112.1; -
DR EMBL; AL109770; CAB524420.1; -
DR GENEDB SPOMBE; SPAC15A10.16; -
DR GO; GO:0005884; C:actin filament; ISS.
DR GO; GO:0016043; P:cell organization and biogenesis; ISS.
DR InterPro; IPR005613; AIP3.
DR Pfam; PF03915; AIP3; 1.
KW Coiled coil; Cytoskeleton.
FT DOMAIN 1009 1096 COILED COIL (POTENTIAL).
SQ SEQUENCE 1385 AA; 154325 MW; ABB3D40CC4FF7537 CRC64;

Query Match 4.2%; Score 159; DB 1; Length 1385;
Best Local Similarity 19.2%; Pred. No. 2;
Matches 147; Conservative 125; Mismatches 293; Indels 202; Gaps 33;

Qy 26 DLNCAQNVVTSSTGDL-----SIPSELENTPESENQYFQSAIWS-----GPIKVKKS 74
Db 326 DRKFSPLKTRTPSLTKSLDPGTSLKSPSLRKSPPS--SPVQKDVYRSNLSRISQANRS 383
Qy 75 DEYTFATSAADNHWVMYDDQEVINKAS-----
Db 384 NVFFCATD---VTRSVSDHRLSSSTINDGEVAPPLPQRSRTISSPNSPISATVLPSST 439
Qy 105 KIRLEKRLYQIKIYOQREN-----PTEKGLDFKL-----YWTDSQNK 143
Db 440 PILLPRGRSSTLSVKNKQFNADGOSTLNSPNSIRETEYAASPKLEADIADDEVTDATSOR 499
Qy 144 EVI-----SSDNLQLPELKQSS---NSRKRSTSGAPTVPPDRDNGI----PDSL 187
Db 500 ELLERQTCKAESSEDTSEISLQKLSLPQVSSSTQIQIOPSSSVPEAASNEIAEKPAVT 559
Qy 188 EVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFKVTGRIDKNV 247
Db 560 AIESIT-----ERKEEAPVISSEKIESGTSISTSTKGLLANPENDSLBELERLIQN- 613
Qy 248 SPEARHPLVAAPIVHVDMENIILSKVEDOSTQNTDSETRTISKNTSTSRTHTSEVHGNA 307
Db 614 NAEQDEFSYKHKYEYSEE-----SGSEDFKSEKDTKGYVISNDDSTQVEESEDKSTP 669
Qy 308 EVHASFFDIGSVSAGFSNNSSTVAID----HSLSLAGERTWAETWGLNTADTARLNAN 363
Db 670 NTGA-----SAKLINDPSSITVSDVVPKPPASPVETEPSSALVSATSPPTNVP 720
Qy 364 I--RYVATGTAPIYVNLPTSLVLGKQTLATIKADENQLSQLAPNNYPSKNLAPIAL 421
Db 721 IVPEAVHLSTA--FSTAPVSTIV-----SNISPLPTVAD--P 753
Qy 422 NAAQDFSTPTTM--NTNQFLELEKT--KQLRLTDQVYGNIAFYNGRVRVYDTGSNW 477
Db 754 NVSGSPSETPISEKPEKVPVVSQTEKALPKPLGVDTKEY--FLRYNNQTRKVKVESPLSN 811
Qy 478 SEVLPTQETTARIIFNGK--DLNLVFERIAANVPSDPLETTKPDMT-LXEALKIARGFN 534
Db 812 ANELGELFSNVYKISFGSDSYELNIED-----PDTKISYLLLEDSDLYKSLVSPMFK 864
Qy 535 EPNGNLQ---YQGDITFDFNFQDQTSQNIKNQLAELNATNIYT-VLDKIKLNAMKNIL 590
Db 865 EQDANKREDPHSGEVSALQHSSAQNTLDDHVNTTTHESPSSAFTEILELKA----- 917
```


